DTIC FILE COPY

1

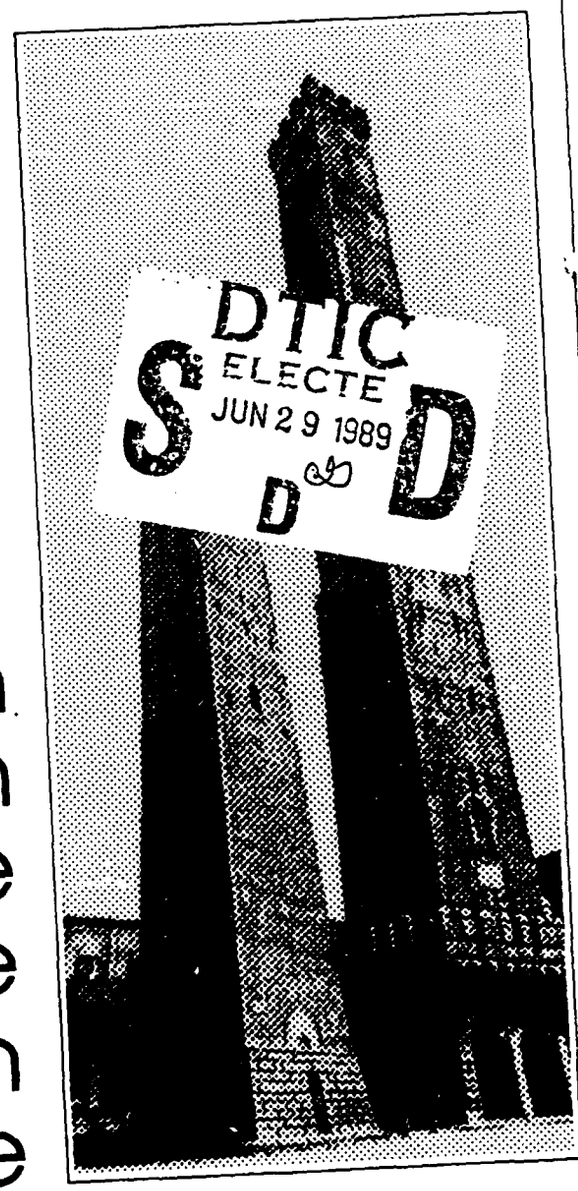
ESSDERC 87

BOLOGNA
ITALY
September 14-17, 1987

AD-A212 842

DTIC
ELECTE
JUN 29 1989
S D
D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited



17th European Solid State Device Research Conference

89 6 29 014

ESSDERC 87

17th European Solid State Device Research Conference

DTIC COPY INSPECTED 6

DAJA45-87-M-0234

Accesion For
NTIS CRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution /
Availability Codes
Dist | Avail and/or Special
A-1

BOLOGNA • ITALY

September 14-17, 1987

A Short History of the University of Bologna

The University of Bologna is the oldest University in the Western world: the birth of the «Studio» was conventionally fixed in 1088, but it had existed long before that: it arose spontaneously in the 11th century as a free association of students and teachers, who founded a school specialized in Roman law, the common cultural heritage of all Western people.

The University of Bologna became famous with the motto «Legum Bononia Mater» and its fame attracted students from all over Italy and Europe.

The students organized themselves into a union called «Universitas Scolarium» to defend their rights: the Union was divided into two: «Universitas Citramontanorum» for italians and «Universitas Ultramontanorum» for foreigners.

The relations between students and the «Comune» of Bologna were sometimes stormy because it was not too easy to reconcile the requests of the «Universitas» with the institutional and social settlement of the city.

By the middle of the 13th century the «Studio» reached its masimum power and began to teach many other subjects: canon law, rhetoric, medicine, philosophy and theology. In the 14th century many other universities in France, England and Spain were founded often with the same statutes and according to the same model of the University of Bologna, known, from then on as "Alma Mater Studiorum".

In 1563 a magnificent building, the Archiginnasio, was built and became the centre of all lectures.

The experimental methods and the new sciences were developed in the Academies, which became the privileged place of scientific and philosophical research.

The most famous of these Academies in Bologna, the Institute of Science, founded museums, libraries and laboratories in the building which became, from the Napoleonic period onwards, the seat of the modern University, transferred here from the Archiginnasio.

Printed in Italy by
TECNOPRINT
Via del Legatore 3, Bologna

Published September, 1987

Welcome to ESSDERC '87

It is a great pleasure and an honour to welcome you to the 17th European Solid State Device Research Conference ESSDERC '87.

This is the first time that the Conference is held in Italy, and we have tried our best to make it a professionally-rewarding as well as an enjoiable meeting. Next november the University of Bologna is going to open its 900th Academic Year, and we are proud to contribute by this scientific event to the celebrations of the "Alma Mater Studiorum".

The general framework of the Conference follows the pattern established over the recent years with a major innovation: the Proceedings inclusive of both invited and extended abstracts of the contributed papers, are made available to all the Delegates attending the Conference and later on will be distributed worldwide by North Holland Publishing Co. We hope this will contribute to the success of the Conference and will increase its impacts on the international scientific community.

Fourteen invited Speakers will present a survey of topics of broad interest in both silicon and compound semiconductors technology and devices, while "special sessions" have been devoted to fast-developping subjects of particular relevance. Furthermore, a one-day Workshop on SOI technologies has jointly been organized with the Electronics Division of the Electrochemical Society, and a panel discussion on "Very High Speed Integrated Circuits" has been sponsored by the XIII Division of the Commission of the European Community.

The Call for Papers resulted in the submission of a record number of 350 papers from 20 Countries, from which about 200 have been selected by the Scientific Program Committee and arranged in three daily parallel sessions and four poster sessions.

The technical program of the Conference appears to be quite demanding and the time allowed to Authors necessarily limited, but we hope that the ESSDERC '87 will be an useful professional experience, an occasion to meet old friends and to make new acquaintances, and an opportunity to discover some of the many beauties of Bologna.

We would like to thank all those who contributed to the organization of the Conference by selecting papers, planning programmes, chairing sessions and doing the many other organizational tasks.

Finally, a special acknowledgement is due to the Organizations, listed in a separate page, who have sponsored and supported the Conference.

Pier Ugo Calzolari
Giovanni Soncini

The Conference has been organized by:

Dipartimento di Elettronica, Informatica e Sistemistica (DEIS)
Università di Bologna
Istituto CNR - LAMEL
Consiglio Nazionale delle Ricerche (CNR)
AEI Gruppo Specialistico CCTE

The organizers gratefully acknowledge support provided to the Conference by:

University of Bologna
Comune di Bologna
Regione Emilia-Romagna
The Regional Board for Economic Dev. (ERVET)
Associazione Elettrotecnica ed Elettronica Italiana (AEI)
IEEE North Italy Section
IBM Italia
SGS Microelettronica

Aurel - Forlì
Bio-Rad - Milan
Datalogic - Bologna
Dynamite Nobel Silicon - Novara
Electrochemical Society - Electronics Division
European Materials Research Society
European Physical Society
Gruppo Nazionale Struttura della Materia (GNSM-CNR)
GTE Telecomunicazioni - Milan
Hewlett Packard Italiana - Milan
Honeywell Bull - Milan
Kormak - Bologna
Officine Galileo - Florence
Olivetti - Ivrea
Selenia - Rome
Siemens - Munich
SMA - Florence
Telettra - Milan
USAF European Office of Aerospace Res. and Dev. - London
US Army European Research Office - London

U.S.A. representative: M. Arienzo, IBM, USA
Japan representative: T. Ikoma, Tokio University
China representative: Tong Qin-Yi, Nanjing Institute of Technology

ORGANIZING COMMITTEE

G. Soncini	Chairman - Istituto CNR - LAMEL Scuola Ingegneria Aerospaziale - Università di Roma «La Sapienza»
P.U. Calzolari	Co-Chairman - Università di Bologna
V.A. Monaco	Finance Chairman - Presidente Gruppo Specialistico AEI-CCTE Chairman IEEE North Italy Section
G.C. Corazza	Presidente Fondazione G. Marconi
D. Nobili	Direttore Istituto CNR-LAMEL
A. Paoletti	Direttore Progetto Finalizzato «Materiali e Dispositivi per l'Elettronica a Stato Solido» del CNR
A. Stella	Presidente GNSM e Direttore CISM
G. Grata	CEE ESPRIT Programme - Head of Microelectronics Division
J.P. Noblanc	Chairman ESSDERC-ESSCIRC Steering Committee
B.L.M. Wilson	Chairman ESSDERC '86
J.P. Nougier	Chairman ESSDERC '88

SCIENTIFIC COMMITTEE

G. Baccarani	Univ. di Bologna	Italy
L. Baldi	SGS Microelettronica	Italy
P. Balk	Aachen Univ.	W. Germany
G. Bentini	CNR-LAMEL	Italy
G. Bomchil	CNET	France
A. Broers	Cambridge Univ.	U.K.
A. Cetronio	SELENIA Roma	Italy
G. Declerck	IMEC	Belgium
V. Ghergia	CSELT Torino	Italy
M. Ilegems	Lausanne Inst. of Technology	Switzerland
F. Klaassen	*PHILIPS*	*Netherlands*
M. Kniepkamp	SIEMENS	W. Germany
H. Martinot	CNRS	France
M. Montier	EFCIS Thomson	France
E. Munoz Merino	Madrid University	Spain
G.F. Piacentini	TELETTRA Milano	Italy
J. Robertson	Edimburg Univ.	U.K.
P. Rocchi	LEP	France
H. Ryssel	Erlangen-Nurnberg University	W. Germany
S. Selberherr	Wien University	Austria
J. Turner	Plessey Research	U.K.
P. Weissglass	Inst. Microwave Technology	Sweden
A. Wieder	SIEMENS	W. Germany
C. Wood	G.E.C. Wembley	U.K.

SPECIAL SESSIONS ORGANIZERS

B. Riccò	Univ. of Bologna	Italy
Monte Carlo simulation for device modelling		
W. Orr-Arienzo	IBM East Fishkill	USA
Gettering phenomena in silicon and related device effects		
L. Treitinger	Siemens	W. Germany
Ultrafast silicon bipolar devices		
M. Melchior	RTH, Zurich	Switzerland
Ultrafast optoelectronics		
J. Robertson	University of Edinburg	U.K.
Test structures for I.C. devices and process evaluation		
F. Fantini	Telettra, Milano	Italy
Reliability Physics		

SUNDAY, SEPTEMBER 13, 1987

CONFERENCE REGISTRATION: 4.00 p.m. - 7.00 p.m. Conference Desk
Palazzo della Cultura e dei Congressi, Fiera District

MONDAY, SEPTEMBER 14, 1987

CONFERENCE REGISTRATION: 8.00 a.m. - 7.00 p.m.

9.30 a.m. - 10.00 a.m.: OPENING CEREMONY - Chairmen: P.U. Calzolari, G. Soncini - Sala Europa

10.00 a.m. - 10.45 a.m.
IP.1: BIPOLAR-CMOS COMBINED PROCESSES: DREAM OR NIGHTMARE?
Invited: P.A.H. Hart - Chairman: W. Heywang - Sala Europa

11.15 a.m. - 12.30 a.m. A1.1: BIPOLAR-CMOS Chairman: G. Zocchi Sala Europa	11.15 a.m. - 11.45 a.m. IP.2: NEW DEVELOPMENTS IN SOLID STATE DETECTORS Invited: E. Gatti - Chairman: A. Paoletti - Sala Italia	
	11.45 a.m. - 12.30 a.m. B1.1: COMPOUND SEMICONDUCTORS TECHNOLOGY I Chairman: H. Martinot - Sala Italia	11.45 a.m. - 12.30 a.m. C1.1: POWER DEVICES I Chairman: A. Stella Sala Bianca
2.00 p.m. - 2.40 p.m. IP.3: PROCESSING AND CHARACTERIZATION OF ULTRASMALL SILICON DEVICES Invited: G. Sai-Halasz - Chairman: F.M. Klaassen - Sala Europa		
2.45 p.m. - 3.30 p.m. A1.2: ULTRASMALL MOS DEVICES Chairman: F.M. Klaassen Sala Europa	2.45 p.m. - 3.45 p.m. B1.2: MONTECARLO DEVICE SIMULATION I Chairman: B. Riccó Sala Italia	2.45 p.m. - 3.45 p.m. C1.2: COMPOUND SEMICONDUCTORS TECHNOLOGY II Chairman: A. Cetronio - Sala Bianca
4.00 p.m. - 5.30 p.m. A1.3: MOS RELIABILITY I Chairman: T.F. Retajczyk Sala Europa	4.00 p.m. - 4.30 p.m. IP.4: VISIBLE LIGHT a-SiC THIN FILM LED AND ITS APPLICATION TO NEW OE FUNCTIONAL ELEMENTS Invited: Y. Hamakawa - Chairman: D. Nobili - Sala Italia	
	4.30 p.m. - 5.30 p.m. B1.3: MONTECARLO DEVICE SIMULATION II Chairman: B. Riccó - Sala Italia	4.30 p.m. - 6.00 p.m. POSTER SESSION P1.1.a: SILICON METALIZATION AND CONTACTS P1.1.b: SENSORS AND DETECTORS Foyer Italia

TUESDAY, SEPTEMBER 15, 1987

9.00 a.m. - 9.40 a.m.
IP.5: TRENDS IN THREE DIMENSIONAL INTEGRATION
Invited: Y. Akasaka - Chairman: R. Van Overstraetene - Sala Europa

9.45 a.m. - 11.00 a.m. B2.1: SOI WORKSHOP I Chairman: D. Mc Caughan Sala Italia	9.45 a.m. - 10.45 a.m. A2.1: ULTRAFAST BIPOLAR I Chairman: L. Treitinger Sala Europa	9.45 a.m. - 11.00 a.m. C2.1: GaAs DEVICE MODELLING Chairman: E. Munoz Merino Sala Bianca	9.45 a.m. - 10.45 a.m. D2.1: GETTERING I Chairman: W. Orr Arienzo Sala Azzurra
11.15 a.m. - 12.45 a.m. B2.2: SOI WORKSHOP II Chairman: D. Mc Caughan Sala Italia	11.15 a.m. - 12.30 a.m. A2.2: ULTRAFAST BIPOLAR II Chairman: L. Treitinger Sala Europa	11.15 a.m. - 11.45 a.m. IP.6: CVD AND INTEGRATED CIRCUITS METALLIZATION Invited: J.O. Carlsson - Chairman: G. Bomchil - Sala Bianca	
		11.45 a.m. - 12.30 a.m. C2.2: MOS MODELLING Chairman: G. Baccarani Sala Bianca	11.45 a.m. - 13.00 a.m. POSTER SESSION P2.1.a: SI PROCESSING AND PROCESS MODELLING P2.1.b: MOS PROCESSING Foyer Italia
2.00 p.m. - 3.45 p.m. B2.3: SOI WORKSHOP III Chairman: D. Mc Caughan Sala Italia	2.00 p.m. - 2.40 p.m. IP.7: BIPOLAR HETEROJUNCTION TRANSISTORS: OUT OF MODELS INTO REALITY Invited: A.J. Holden - Chairman: M. Kniepkamp - Sala Europa		
	2.45 p.m. - 3.30 p.m. A2.3: ULTRAFAST BIPOLAR III Chairman: L. Treitinger Sala Europa	2.45 p.m. - 3.30 p.m. C2.3: MOS MODELLING II Chairman: S. Selberherr Sala Bianca	2.45 p.m. - 3.30 p.m. D2.2: GETTERING II Chairman: W. Orr-Arienzo Sala Azzurra
4.00 p.m. - 5.30 p.m. B2.4: GaAs MESFET RELIABILITY Chairman: F. Fantini Sala Italia	4.00 p.m. - 5.45 p.m. A2.4: CMOS TECHNOLOGY Chairman: L. Baldi Sala Europa	4.00 p.m. - 4.30 p.m. IP.8: ION BEAM LITHOGRAPHY Invited: H. Löschner - Chairman: A. Broers - Sala Bianca	
		4.30 p.m. - 5.45 p.m. C2.4: POWER DEVICES II Chairman: E. Rimini Sala Bianca	4.30 p.m. - 6.00 p.m. POSTER SESSION P2.2.a: MOS DEVICES, MEASUREMENTS AND SPECIAL DEVICES P2.2.b: MOS RELIABILITY II Foyer Italia

WEDNESDAY, SEPTEMBER 16, 1987

9.00 a.m. - 9.40 a.m.
IP.9: TRENDS IN NON-VOLATILE MEMORY DEVICES AND TECHNOLOGIES
Invited: H.E. Maes - Chairman: G. Declerck - Sala Europa

9.45 a.m. - 10.45 a.m. A3.1: MOS MEMORIES Chairman: G. Declerck Sala Europa	9.45 a.m. - 10.45 a.m. B3.1: LATCH-UP Chairman: M. Montier Sala Italia	9.45 a.m. - 10.45 a.m. PANEL SESSION VERY HIGH SPEED INTEGRATED CIRCUITS Chairman: H. Wieder Sala Bianca	9.45 a.m. - 10.45 a.m. D3.1: LASER I Chairman: C. Wood Sala Azzurra
11.15 a.m. - 12.30 a.m. A3.2: THIN MOS DIELECTRICS Chairman: P. Balk Sala Europa	11.15 a.m. - 11.45 a.m. IP.10: LIGHT-GUIDED ETCHING FOR III-V SEMICONDUCTOR DEVICE FABBRICATION Invited: D. Podlesnik - Chairman: G. Bentini Sala Italia	11.15 a.m. - 11.45 a.m. IP.11: DESIGN AND PERFORMANCE OF HIGH POWER SEMICONDUCTOR LASERS Invited: K. Mettler - Chairman: T. Ikegami Sala Bianca	
	11.45 a.m. - 12.30 a.m. B3.2: COMPOUND SEMICONDUCTORS TECHNOLOGY III Chairman: G.F. Piacentini Sala Italia	11.45 a.m. - 12.30 a.m. PANEL SESSION VERY HIGH SPEED INTEGRATED CIRCUITS Chairman: H. Wieder Sala Bianca	11.45 a.m. - 12.30 a.m. D3.2: LASER II Chairman: J. Turner Sala Azzurra

2.00 p.m. - 7.00 p.m.
EXCURSION BY COACH TO RAVENNA

9.00 p.m. - 12.00 p.m.
GALA DINNER IN "PALAZZO ALBERGATI"
Bus departure to Palazzo Albergati: from Palazzo dei Congressi: at 8.30 p.m.
from Central Railway Station: at 8.30 p.m.
from Piazza Maggiore: at 8.30 p.m.

THURSDAY, SEPTEMBER 17, 1987

9.00 a.m. - 9.40 a.m.
IP.12: INTEGRATED OPTICS: $LiNbO_3$ OR SEMICONDUCTORS?
Invited: M. Papuchon - Chairman: J.P. Noblanc - Sala Europa

9.45 a.m. - 10.45 a.m. A4.1: INTEGRATED OPTOELECTRONICS I Chairman: P. Rocchi Sala Europa	9.45 a.m. - 10.45 a.m. B4.1: BIPOLAR MODELLING I Chairman: H.C. De Graaff Sala Italia	9.45 a.m. - 10.45 a.m. C4.1: TEST CHIPS Chairman: J. Robertson Sala Bianca
11.15 a.m. - 12.30 a.m. A4.2: INTEGRATED OPTOELECTRONICS II Chairman: V. Ghergia Sala Europa	11.15 a.m. - 11.45 a.m. IP.13: THE PHYSICS OF SILICIDE BASE TRANSISTORS Invited: E. Rosencher - Chairman: H. Ryssel Sala Italia	
	11.45 a.m. - 12.30 a.m. B4.2: BIPOLAR MODELLING II Chairman: H.C. De Graaff Sala Italia	11.45 a.m. - 13.00 a.m. POSTER SESSION P4.1.a: GaAs TECHNOLOGY AND DEVICES P4.1.b: OPTOELECTRONICS Foyer Italia

2.00 p.m. - 2.40 p.m.
IP.14: HEMT AND MQW BASED IC. TECHNOLOGY FOR ULTRA-HIGH SPEED SIGNAL PROCESSING
Invited: A. Christoy - Chairman: B.L.M. Wilson - Sala Europa

2.45 p.m. - 4.00 p.m. A4.3: BIPOLAR TECHNOLOGY Chairman: A. Wieder Sala Europa	2.45 p.m. - 4.00 p.m. B4.3: ULTRAFAST OPTOELECTRONICS Chairman: M. Melchior Sala Italia	2.45 p.m. - 4.00 p.m. C4.2: LATE NEWS Chairman: P.U. Calzolari Sala Bianca

Technical Program Contents

MONDAY, SEPTEMBER 14, 1987

SALA EUROPA - CHAIRMEN: P.U. CALZOLARI AND G. SONCINI

9.30 *Opening Ceremony*

SALA EUROPA - CHAIRMAN: W. HEYWANG

10.00 IP.1 ***Bipolar-CMOS combined processes. Dream or nightmare? (invited)*** 1
P.A.H. Hart
Philips Research Labs., Eindhoven, NL

SALA ITALIA - CHAIRMAN: A. PAOLETTI

11.15 IP.2 ***New developments in solid-state detectors (invited)*** 9
E. Gatti, A. Longoni and M. Sampietro
Politecnico of Milan, Milan, Italy

SESSION A1.1: BIPOLAR-CMOS
Monday, September 14, 1987
SALA EUROPA - CHAIRMAN: G. ZOCCHI

11.15 **A1.1.1 The usefulness of advanced drain structures as emitters in scaled BICMOS** 25
J. Winnerl, F. Neppl, B. Vollmer, M. Stegherr and B. Pfäffel
Siemens, München, FRG

11.30 **A1.1.2 1.2 μm Bi-CMOS technology with high performance ECL** 29
H. Iwai, Y. Niitsu, G. Sasaki, M. Norishima, K. Shino, Y. Unno, K. Tsugaru, H. Hara, Y. Sugimoto and K. Kanzaki
Toshiba, Kawasaki, Japan

11.45 **A1.1.3 Multipower BCD 250V: a versatile technology to realize high performance PICs** 33
A. Andreini, C. Contiero and P. Galbiati
SGS Microelettronica, Milan, Italy

12.00 **A1.1.4 A 1.5 μm analogue CMOS process**
J.N. Ellis, J.G. Daniels and D.J. Wilcox
Plessey, Caswell, UK
(extended abstract not available at the time of printing)

12.15 **A1.1.5 Rapid thermal processing of polysilicon emitter bipolar transistors in a combined CMOS/Bipolar process** 37
L.A. Grant, D.W. McNeill
STC Technology, Harlow, UK
P.F. Blomley
LSI Logic, Bracknell, UK

SESSION B1.1: COMPOUND SEMICONDUCTORS TECHNOLOGY I
Monday, September 14, 1987
SALA ITALIA - CHAIRMAN: H. MARTINOT

11.45 **B1.1.1 Substrate influences on the activation of ion implanted Si in GaAs** 43
R.D. Schnell and H. Schink
Siemens, München, FRG

12.00 **B1.1.2 Comparison of rapid annealing and furnace annealing of Si implanted into GaInAs** 47
J. Splettstösser, H. Heesel, U. Breuer, W. Albrecht and H. Beneking
Aachen Technical Univ., Aachen, FRG
D. Schmitz
Aixtron, Aachen, FRG
J. Selders
Telefunken Electronic, Heilbronn, FRG

12.15 **B1.1.3 Yield-performance considerations for ion-implanted GaAs integrated circuits based on substrate material properties** 51
C. Lanzieri, R. Graffitti, C. Calori, S. Rapisarda and A. Cetronio
SELENIA, Rome, Italy

SESSION C1.1: POWER DEVICES I
Monday, September 14, 1987
SALA BIANCA - CHAIRMAN: A. STELLA

11.45 **C1.1.1 A novel 1500 volt IGBT device with improved turn-off performance** 57
J. Blake, H.E. Brockman and R.W. Cooper
Philips Res. Labs, Redhill, UK

12.00 **C1.1.2 Design of a 1600 V power bipolar mode FET** 61
P. Spirito and G. Vitale
Naples Univ., Dptm. of Electronics Engineering, Naples, Italy
G. Busatto
IRECE-CNR, Naples, Italy
G. Ferla and S. Musumeci
SGS Microelettronica, Catania, Italy

12.15 **C1.1.3 Modelling bipolar transitor second breakdown during turn-off by solution of the fundamental device equations** 65
S.A. Higgins, M.K. Johnson, P.A. Gough and J.A.G. Slatter
Philips Res. Labs., Redhill, UK

SALA EUROPA - CHAIRMAN: F.M. KLAASSEN

14.00 **IP.3** ***Processing and characterization of ultrasmall silicon devices (invited)*** 71
G.A. Sai-Halasz
IMB, T.J. Watson Res. Center, Yorktown Heights, NY, USA

SESSION A1.2: ULTRASMALL MOS DEVICES
Monday, September 14, 1987
SALA EUROPA - CHAIRMAN: F.M. KLAASSEN

14.45 **A1.2.1 Half-micrometer N-MOS technology using X-Ray lithography** 83
V. Lauer, F. Bauer, J. Korec and P. Balk
Aachen Technical Univ., Aachen, FRG
H. Huber
Fraunhofer-Institut, Berlin, FRG

15.00 **A1.2.2 Offset diffused drain transitors for half-micron CMOS** 87
P.H. Woerlee, C.A.H. Juffermans, H. Lifka, F.M. Oude Lansink, H.J.H. Merks-Eppingbroek, T. Poorter and A.J. Walker
Philips Res. Labs. Eindhoven, NL

15.15 **A1.2.3 0.5 μm CMOS Device design and characterization** 91
A.I. Hanafi, M.R. Wordeman, L.K. Wang, Y. Taur, J.Y.C. Sun, R.H. Dennard, D.S. Zicherman and M.D. Rodriguez
IBM, T.J. Watson Res. Center, Yorktown Heights, NY, USA
N. Haddad, A. Edenfeld and M. Polavarapu
IBM FSD, Manassas, VA, USA

SESSION B1.2: MONTE CARLO DEVICE SIMULATION I
Monday, September 14, 1987
SALA ITALIA - CHAIRMAN: B. RICCO'

14.45 **B1.2.1 Monte Carlo simulation of semiconductor devices: a critical review** 97
P. Lugli and C. Jacoboni
Modena Univ., Modena, Italy

15.00 **B1.2.2 A Monte Carlo study of diffusion coefficients of two-dimensional electron gas in HEMT AlGaAs-GaAs structures** 103
J. Zimmerman and Y. Wu
Lille Univ., Villeneuve d'Ascq, France

15.15 **B1.2.3 Monte Carlo simulation of classical and inverted MODFET structures** 107
R. Fauquembergue, M. Pernisek, J.L. Thobel and P. Bourel
Lille Univ., Villeneuve d'Ascq, France

15.30 **B1.2.4 The particle simulation of self-aligned GaAs MESFETs with a sub-micrometer gate-lenght** 111
Y. Yamada, S. Ikeda and N. Shimojoh
Kumamoto Univ., Kumamoto, Japan

SESSION C1.2: COMPOUND SEMICONDUCTORS TECHNOLOGY II
Monday, September 14, 1987
SALA BIANCA - CHAIRMAN: A. CETRONIO

14.45 **C1.2.1 Redistribution of ion-implanted mercury during rapid thermal annealing of GaInAs and InP** 117
J.H. Wilkie and B.J. Sealy
Surrey Univ., Guildford, UK

15.00 **C1.2.2 Rapid thermal annealing of Be implants into undoped InP** 121
W. Häussler,
Siemens, München, FRG

15.15 **C1.2.3 Accumulation of implanted hydrogen at the superlattice/substrate interface** 125
J.M. Zavada
USARDSG, London, UK
R.G. Wilson
Hughes Res. Labs, Malibu, CA, USA
S.W. Novak
Evans Ass., Redwood City, CA, USA

15.30 **C1.2.4 Multipolar plasma treatments of InGaAs surface for MIS devices aplications** 129
M. Renaud, P. Boher, J. Barrier, J. Schneider and J.P. Chané
LEP, Limeil Brevannes, France

SALA ITALIA - CHAIRMAN: D. NOBILI

16.00 **IP.4** ***Visible light a-SiC thin film LED and its application to new OE-functional elements*** 135
Y. Hamakawa, D. Krungam, H. Okamoto and H. Takakura
Osaka Univ., Osaka, Japan

SESSION A1.3: MOS RELIABILITY I
Monday, September 14, 1987
SALA EUROPA - CHAIRMAN: T.F. RETAJCZYK

16.00 **A1.3.1 Comparison between hot-carrier drift and radiation damage in MOS devices** 145
A.G. Sabnis
AT & T Bell Labs., Allentown, PA., USA

16.15 **A1.3.2 The role of holes and electrons in the aging of MOS transistors** 151
M. Tosi, L. Baldi and F. Maggioni
SGS Microelettronica, Milan, Italy

16.30 **A1.3.3 The voltage dependence of degradation in N-MOS transistors** 155
B.S. Doyle, M. Bourcerie, J.C. Marchetaux and A. Boudou
BULL S.A., Les Clayes sous Bois, France

16.45 **A1.3.4 Characterisation and analysis of hot-carrier degradation in p-channel transistors using constant current stress experiments** 159
R. Bellens, P. Heremans, G. Groeseneken and H.E. Maes
IMEC, Leuven, Belgium

17.00 **A1.3.5 Correlation between flatband voltage shift in MOS capacitors and endurance degradation of EEPROM cells** 163
J. Manthey, M. Dutoit and M. Ilegems
Federal Institute of Technology, Lausanne, Switzerland

17.15 **A1.3.6 Degradation phenomena of tunnel oxide floating gate EEPROM devices** 167
J.S. Witters, G. Groeseneken and H.E. Maes
IMEC, Leuven, Belgium

SESSION B1.3: MONTE CARLO DEVICE SIMULATION II
Monday, September 14, 1987
SALA ITALIA - CHAIRMAN: B. RICCO'

16.30 **B1.3.1 A Monte Carlo approach to the study of the dirft-diffusion transport model** 173
C. Mantilli, F. Venturi, B. Riccò and E. Sangiorgi
Bologna Univ., Bologna, Italy

16.45 **B1.3.2 Hot electron dynamics Monte Carlo simulation in heterostructure semiconductor devices** 177
F. Antonelli
IBM, Roma, Italy
P. Lugli
Modena Univ., Modena, Italy

17.00 **B1.3.3 A Monte Carlo analysis of diffusion-noise properties in GaAs-AlGaAs Quantum Wells** 181
S.M. Goodnick
Oregon State Univ., Corvallis, Oregon, USA
R. Brunetti
Modena Univ., Modena, Italy

17.15 **B1.3.4 A deterministic particle method for the semiconductor Boltzmann equation** 185
P. Degond, B. Niclot and F. Guyot
Ecole Polytechnique, Palaiseau, France

SESSION P1.1: POSTERS

16.30 **Monday, September 14, 1987**

a. Silicon Metalization and contacts

P1.1.1 TiW as a direct contact material and a diffusion barrier to n^+ and p^+ implanted areas 191
A Lindberg and M. Östling
Inst. of Microwave Technology, Stockholm, Sweden
H. Norström and U. Wennström
RIFA, Stockholm, Sweden

P1.1.2 Comparison of the behaviour of As during Ti and WSi_2 formation 197
J. Torres, J.C. Oberlin, G. Bomchil and A. Perio
CNET, Meylan, France
D. Levy
Bull, Les Clayes-sous-Bois, France
A. Saulnier, J.P. Ponpon and R. Stuck
CNRS, Centre de Recherches Nucleaires, Strasbourg, France

P1.1.3 LPCVD tungsten filled vias for multilayer metalization 201
S.L. Zhang, R. Buchta and T. Johansson
Inst. of Microwave Technology, Stockholm, Sweden
H. Norström and U. Wennström
RIFA, Stockholm, Sweden

P1.1.4 Al/TiN/$TiSi_2$ contacts to ultra shallow junctions 205
E. Ling, H.S. Gamble, B.M. Armstrong and J.H. Montgomery
The Queen's Univ., Belfast, N. Ireland

P1.1.5 The effect of ion-irradiation and rapid thermal annealing on $TiSi_2$ and $MoSi_2$ 209
L. Grönberg, J. Saarilahti and I. Suni
Techn. Res. Centre, Otakaari, Finland
Ch. Krontiras
Patras Univ, Patras, Greece

P1.1.6 Evaluation of electromigration activation energy by means of noise measurements and MTF tests 213
A. Diligenti, P.E. Bagnoli and B. Neri
Pisa Univ., Pisa, Italy
G. Specchiulli
Telettra, Milan, Italy

P1.1.7 Electromigration in narrow Al(Si) stripes for VLSI: comparison between MTF and resistometric methods for life prediction 217
G. Specchiulli and F. Fantini
Telettra, Milan, Italy
G. De Santi
SGS Microelettronica, Milan, Italy

P1.1.8 Electromigration control: the accelerated BEM test 221
L. Bacci, C. Caprile and G. De Santi
SGS Microelettronica, Milan, Italy

P1.1.9 Electrical and structural characterization of electromigration in Al-Si/Ti multilayer interconnects 225
M. Finetti, A. Scorzoni, A. Armigliato and A. Garulli
CNR-LAMEL, Bologna, Italy
I. Suni
Technical Res. Center, Otakaari, Finland

P1.1.10 Plasma-enhanced chemical vapour deposition on silicon and silicon dioxide substrates 229
C.M.T. Hodson, J. Wood and M. Middleton
York Univ., Dptm. of Electronics, York, UK

P1.1.11 Plasma anodization of silicides 233
B. Pelloie, J. Perriere, J.P. Enard and A. Laurent
Univ. of Paris VII, Paris, France
I. Montero, A. Climent, R. Perez and J.M. Martinez-Duart
Univ. Autonoma Cantoblanco, Madrid, Spain
R. Nipoti and S. Guerri
CNR-LAMEL, Bologna, Italy

P1.1.12 A novel process for silicide formation using dynamic recoil mixing 237
L.I. Haworth, R. Holwill and J.M. Robertson
Edimburgh Univ., Edimburgh, UK
A.E. Hill and R. Pilkington
Univ. of Salford, Salford, UK

b. Sensors and detectors

P1.1.13 Simulation of gate-controlled double-injection SOI structures: application to micromagnetodiode sensors 241
G. Dimopoulos, F. Balestra, A. Chovet, M. Benachir and J. Brini
ENSERG, Grenoble, France

P1.1.14 Optimization of p-implanted silicon bolometers 245
E. Baciocco, C. Boragno and U. Valbusa
Dip. Fisica, Univ. of Genova, Italy
C. Bresolin and G. Pignatel
SGS Microelettronica, Milan, Italy

P1.1.15 Large-area silicon detectors for calorimetry in high-energy physics 249
B. Passerini, M. Zambelli and P.E. Zani
Ansaldo, Genova, Italy

P1.1.16 A Magnetic field sensor using a graded gate potential 253
B.S. Gill and E.L. Heasell
Univ. of Waterloo, Waterloo, Ontario, Canada

P1.1.17 Numerical modelling of magnetic field sensitive MOSFET 257
Ho Yie, Wei Tongli and Shen Kechang
Nanjing Inst. of Technology, Nanjing, China

P1.1.18 Enzymatic reaction heat detection by a pyroelectric device
R. Mercuri
I.E.S.S.-CNR, Roma, Italy

A. Barbaro, V. Baroncelli, C. Colapicchioni, R. Colilli, I. Giannini and M. Lupoli
Eniricerche, Monterotondo (Roma), Italy
A. D'Amico
L'Aquila Univ., L'Aquila, Italy
(extended abstract not available at the time of printing)

P1.1.19 Research on the growth condition of CdTe single crystal for y-ray detectors 261
K. Mochizuki and K. Masumoto
Tohoku Univ., Sendai, Japan

TUESDAY, SEPTEMBER 15, 1987

SALA EUROPA - CHAIRMAN: R. VAN OVERSTRAETEN

9.00 **IP.5** ***Trends in three-dimensional integration (invited)*** 265
Y. Akasaka
LSI Lab., Mitsubishi Electric Co., Japan

SESSION A2.1: ULTRAFAST BIPOLAR I
Tuesday, September 15, 1987
SALA EUROPA - CHAIRMAN: L. TREITINGER

9.45 **A2.1.1 History, present trends and scaling of silicon bipolar tecnology *(invited)*** 277
Tak H. Ning
IBM, T.J. Watson Res. Center, Yorktown Heights, NY, USA

10.15 **A2.1.2 PABLO vs. double-poly. A Comparison of two high-performance bipolar technologies** 283
R.A. Van Es and D.J.W. Noorlag
Philips Research Labs., Eindhoven, NL

10.30 **A2.1.3 Rapid annealing for shallow junction formation** 287
A.E. Michel
IBM, T.J. Watson Res. Center, Yorktown Heights, NY, USA

SESSION B2.1: SOI WORKSHOP I
Tuesday, September 15, 1987
SALA ITALIA - CHAIRMAN: D. MC CAUGHAN

9.45 **B2.1.1 Recrystallization of SOI films by a pseudoline electron beam *(invited)*** 293
H. Ishiwara and S. Horita
Tokio Inst. of Technology,Yokohama, Japan

10.15 **B2.1.2 A comparative study of starting materials for SOI device fabrication obtained by different laser recrystallization procedures and material structures** 299
D.J. Wouters, M.R. Tack, P.W. Mertens,, H.E. Maes and C.L. Claeys
IMEC, Leuven, Belgium

10.30 **B2.1.3 A 3D-SOI intelligent power structures** 303
B. Dunne, C.G. Cahill, A. Mathewson and W.A. Lane
University Collegé Cork, Ireland
M. Montier and D. Chapuis
Thompson EFCIS, France

10.45 **B2.1.4 Analysis of nonuniformly doped SOI MOSFET's** 307
P. Paelinck, O. Vancauwenberghe and F. Van de Wiele
Univ. Catholique de Louvain, Louvain-La-Neuve, Beglium

SESSION C2.1: GaAs DEVICE MODELLING
Tuesday, September 15, 1987
SALA BIANCA - CHAIRMAN: E. MUNOZ MERINO

9.45 **C2.1.1 Modelling of the drain lag effect in GaAs MESFET's and its impact on digital IC's** 313
T. Ducourant and M. Rocchi
LEP, Limeil-Brevannes, France

10.00 **C2.1.2 Modelling and simulation of wave propagation effects in MESFET devices based on physical models** 317
G. Ghione and C.U. Naldi
Politecnico of Torino, Dept. of Electronics, Torino, Italy

10.15 **C2.1.3 Accurate nonlinear characterization and modelling of the GaAs FET** 321
Y. Bonnaire and E. Allamando
Centre Hyperfrequencies et Semiconducteurs
Lille Univ., Villeneuve D'Ascq, France

10.30 **C2.1.4 A CAD model for heterojunction bipolar transistors**
J. Dangla, M. Laporte, P.Y. Merrer and E. Caquot
CNET, Bagneux, France
(extended abstract not available at the time of printing)

10.45 **C2.1.5 A three-dimensional model with distributed elements for GaAs MESFETs and similar devices** 325
W. Wiesbeck, S. Haffa and H.P. Feldle
Karlsruhe Univ., Karlsruhe, FRG

SESSION D2.1: GETTERING I
Tuesday, September 15, 1987
SALA AZZURRA - CHAIRMAN: W. ORR - ARIENZO

9.45 **D2.1.1 Instrinsic gettering: sense or nonsense?** 331
J. Vanhellemont and C. Claeys
IMEC, Leuven, Belgium

10.00 **D2.1.2 Lifetime engineering by oxygen precipitation in silicon** 335
M. L. Polignano and G.F. Cerofolini
SGS Microelettronica, Milan, Italy
H. Bender and C. Claeys
IMEC, Heverlee, Belgium
J. Reffle
Wacker Chemitronic, Burghausen, FRG

10.15 **D2.1.3 A model for oxygen phase transition kinetics in CZ-grown silicon and its application to IC processes** 339
M. Pagani
Dynamit Nobel Silicon, Novara, Italy
W. Huber
Dynamit Nobel Silicon Tech. Center, Sunnyvale, CA. USA

10.30 **D2.1.4 Influence of intrinsic gettering on silicon recombination properties and their relation to device performance** 343
M. Kittler, H. Richter and W. Seifert
Academy of Sciences, Frankfurt (Oder), DDR

SALA BIANCA - CHAIRMAN: G. BOMCHIL

11.15 **IP.6** ***CVD and integrated circuits metallization (invited)*** 347
J.O. Carlsson
Uppsala Univ. Institute of Chemistry, Uppsala, Sweden

SESSION A2.2: ULTRAFAST BIPOLAR II
Tuesday, September 15, 1987
SALA EUROPA - CHAIRMAN: L. TREITINGER

11.15 **A2.2.1 Self aligned technology for sub-100 nm deep base junction transistors** 361
M. Nakamae
NEC, Kanagawa, Japan

11.30 **A2.2.2 Vertical scaling considerations for polysilicon-emitter bipolar transistors** 365
H. Schaber, J. Bieger, B. Benna and T. Meister
Siemens, München, FRG

11.45 **A2.2.3 Trench isolation schemes for bipolar devices: benefits and limiting aspects** 369
H. Goto and K. Inayoshi
Fujitsu, Kawasaki, Japan

12.00 **A2.2.4 A salicide base contact technology (SCOT) for use in high speed bipolar VLSI** 373
T. Hirao, T. Ikeda and Y. Kuramisu
LSI Lab., Mitsubishi, Itami, Japan

12.15 **A2.2.5 Trends in heterojunction silicon bipolar transistors** 377
R. Mertens, J. Nijs, J. Symons, K. Baert and M. Ghannam
IMEC, Leuven, Belgium

SESSION B2.2: SOI WORKSHOP II
Tuesday, September 15, 1987
SALA ITALIA - CHAIRMAN: D. MC CAUGHAN

11.15 **B2.2.1 Electrical parameters of SOI material obtained by ZMR and oxidized porous silicon *(invited)*** 385
M. Haond, G. Bomchil, J.L. Regolini, D. Bensahel, D. Dutartre, D.P. Vu, K. Barla, H. Halimaoui, R. Herino
CNET, Meylan, France
A. Monroy, S. Thouret and Y. Gris
Thomson, France

11.45 **B2.2.2 Porous anodised silicon for full dielectric isolation: the development of an $n/n^+/n$ device route** 391
D. Brumhead, J.G. Castledine and J.M. Keen
RSRE, Great Malvern, UK
J.M. Cole, L.G. Earwaker, J.P.G. Farr, P.E. Grzeszczyk, J. L'Ecuyer, M.H. Loretto and I.M. Sturland
Birmingham Univ., Birmingham, UK

12.00 **B2.2.3 STACMOS: a basic 3-dimensional CMOS process** 395
R. Buchner, K. Haberger, P. Seegebrecht and P. Panish
Fraunhofer-Inst., München, FRG

12.15 **B2.2.4 Volume inversion in SOI MOSFET's with double gate control: a new transistor operation with greatly enhanced performance** 399
F. Balestra, S. Cristoloveanu, M. Benachir, J. Brini and T. Elewa
ENSERG, Grenoble, France

12.30 **B2.2.5 A new type of high performance device for VLSI digital system** 403
Xu Xiao-Li, Tong Qin-Yi and Xong He-Ming
Microelectronics Center, Nanjing Inst. of Technology, Nanjing, China

SESSION C2.2: MOS MODELLING I
Tuesday, September 15, 1987
SALA BIANCA - CHAIRMAN: G. BACCARANI

11.45 **C2.2.1 Comparison of long-and short-channel MOSFET's carried out by 3D-minimos** 409
M. Thurner and S. Selberherr
Technische Universität, Wien, Austria

12.00 **C2.2.2 Three-dimensional simulation of a narrow-width MOSFET** 413
P. Ciampolini, A. Gnudi, R. Guerrieri, M. Rudan and G. Baccarani
Bologna Univ., Bologna, Italy

12.15 **C2.2.3 2-D and 3-D capacitance effects in MOS VLSI** 417
J.H.M. Quint, F.M. Klaassen and R. Petterson
Philips Research Labs., Eindhoven, NL

11.45

SESSION P2.1: POSTERS
Tuesday, September 15, 1987

a. Si processing and process modelling

P2.1.1 Implantation and diffusion modelling of boron in silicon 423
An De Keersgieter, L. Dupas and K. De Meyer
IMEC, Leuven, Belgium

P2.1.2 Shallow junctions of boron implanted in Ge^+ preamorphized <100> Si wafers 429
A. La Ferla, V. Raineri and E. Rimini
Deptm. of Physics, Catania Univ., Italy
S. Cannavò and G. Ferla
SGS Microelettronica, Catania, Italy

P2.1.3 The effect of high pressure steam oxidation on phosohorous diffusion in silicon 433
Wu Bai-Lu, Xue Shi-Ying
Academia Sinica, China
Zhang Ai-Zhen
Beijing Inst. of Semic. Devices, China
Li Shy-Lin
Nat. Inst. of Metrology, China

P2.1.4 Shallow junction formation using $COSi_2$ as a diffusion source 437
V. Probst and H. Schaber
Siemens, München, FRG
P. Lippens, L. Van den Hove, K. Maex and R. De Keersmaecker
IMEC, Leuven, Belgium

P2.1.5 2-D effects during isolation process - Experiments and Simulation 441
A. Seidl
Inst. für Festkörpertechnologie, München, FRG
V. Huber
Siemens, München, FRG

P2.1.6 Verification of ion implantation models by Monte Carlo simulations 445
G. Hobler and S. Selberherr
Wien Technical Univ., Wien, Austria

P2.1.7 A simplified model for the characterization of antimony ion implantation and diffusion in silicon 449
R.I. Fonseca
Escola Politecnica de USP, Sao Paulo, Brasil

P2.1.8 Glass reflow modeling for process optimization 453
A. Tissier, A. Poncet and J.F. Teissier
CNET, Grenoble, France

P2.1.9 Monte-Carlo ion implanation and COMPOSITE 457
A. Barthel, J. Lorenz and H. Ryssel
Fraunhofer-Arbeitsgruppe für Integrierte Schaltungen, Erlangen, FRG
H. Ryssel
Universität Erlangen- Nürnberg, FRG

P2.1.10 Equilibrium solubility of arsenic and antimony in silicon 461
R. Angelucci, A. Armigliato, E. Landi, D. Nobili and S. Solmi
CNR-LAMEL, Bologna, Italy

P2.1.11 Diffusion and solubility of gold implanted in silicon 465
S. Coffa, L. Calcagno and S.U. Campisano
Dip. Fisica, Catania Univ., Catania, Italy
G. Calleri and G. Ferla
SGS Microelettronica, Catania, Italy

P2.1.12 Open stencil masks for ion projection lithography 469
L.M. Buchmann, L. Csepregi and K.P. Müller
Fraunhofer-Inst. für Mikrostrukturtechnik, Berlin, FRG

b. MOS processing

P2.1.13 A new isolation process for VLSI devices 473
E. Figueras, J.L. Coppee and F. Van de Wiele
Univ. Catholique, Louvain-La-Neuve, Belgium

P2.1.14 Oxidation induced local channel dopant accumulation 477
C. Mazuré and M. Orlowski
Siemens, München, FRG

P2.1.15 Redistribution of flourine from BF_2^+ implants in MOS structures 481
H.J. Whitlow, C. Zaring and C.S. Petersson
The Royal Inst. of Technology, Stockholm, Sweden
J. Keinonen
Helsinki Univ. Accelerator Laboratory, Helsinki, Finland

P2.1.16 Dopant redistribution from ion implanted WSi_2 on poly-Si 483
S. Nygren, D. Levy, G. Goltz and J. Torres
CNET/CNS, Meylan, France

SALA EUROPA - CHAIRMAN: M. KNIEPKAMP

14.00 **IP.7** ***Bipolar heterojunction transistors: out of models into reality (invited)*** 487
A.J. Holden
Plessey Research, Caswell, UK

SESSION A2.3: ULTRAFAST BIPOLAR III
Tuesday, September 15, 1987
SALA EUROPA - CHAIRMAN: L. TREITINGER

14.45 **A2.3.1 Characterization and optimization of bipolar technologies by means of high speed circuit design** 499
W. Wilhelm
Siemens, München, FRG

15.00 **A2.3.2 Physical modelling problems of ultrafast silicon bipolar transistors** 503
H.C. de Graaf and G.A.M. Hurkx
Philips Research Labs., Eindhoven, NL

15.15 **A2.3.3 An advanced bipolar process using trench isolation and polysilicon emitter for high speed VLSI** 507
M. Roche, G. Borel, Jl. Imbert, D. Thomas, L. Fritsch and D. Celi
Thompson, St.-Egrève, France
P. Hunt
Plessey, Caswell, UK
A. Hefner
Telefunken, Heilbronn, FRG

SESSION B2.3: SOI WORKSHOP III
Tusday, September 15, 1987
SALA ITALIA - CHAIRMAN: D. MC CAUGHAN

14.00 **B2.3.1 SOI structures by ion implantation and annealing in a temperature gradient *(invited)*** 513
G.K. Celler
ATT Bell Labs, Murray Hill, N.J., USA

14.30 **B2.3.2 Oxygen implantation for SOI device technologies** 521
K. Yallup
Analog Devices, Leuven, Belgium
An De Veirman
Antwerp Univ., Antwerp, Belgium
L. Dupas and K. De Meyer
IMEC, Leuven, Belgium

14.45 **B2.3.3 Small geometry SOI/CMOS devices on SIMOX substrates** 525
J.R. Davis
British Telecom, Ipswich, UK
K. Reeson and P.L.F. Hemment
Univ. of Surrey, Guildford, UK

15.00 **B2.3.4 Carrier generation and trapping properties in SIMOX structures** 529
T. Elewa, H. Haddara and S. Cristoloveanu
ENSERG, Grenoble, France

15.15 **B2.3.5 Small geometry NMOS transistors in silicon-on sapphire using rapid annealed source/drains and improved crystalline quality silicon films** 533
M. Field, N.E.B. Cowern and D.J. Godfrey
GEC, Hirst Res. Centre, Wembley, UK

15.30 **B2.3.6 Silicon I.C. technology using complementary MESFET's** 537
P.A. Tove, K.E. Bohlin, H. Norde, U. Magnusson, J. Tirén, A. Söderbärg, M. Rosling, F. Masszi and J. Nylander
Inst. of Technology, Uppsala Univ., Uppsala, Sweden

SESSION C2.3: MOS MODELLING II
Tuesday, September 15, 1987
SALA BIANCA - CHAIRMAN: S. SELBERHERR

14.45 **C2.3.1 The voltage-doping transformation: a new approach to the modelling of MOSFET short-channel effects** 543
T. Skotnicki, G. Merckel and T. Pedron
CNET-CNS, Meylan, France

15.00 **C2.3.2 A novel realistic model for threshold voltage of short channel MOSFET's** 547
M. Orlowski and Ch. Werner
Siemens, München, FRG

15.15 **C2.3.3 Sensitivity analysis for device design** 551
A. Gnudi, P. Ciampolini, R. Guerrieri, M. Rudan and G. Baccarini
Bologna Univ., Bologna, Italy

SESSION D2.2: GETTERING II
Tuesday, Spetember 15, 1987
SALA AZZURRA - CHAIRMAN: W. ORR-ARIENZO

14.45 **D2.2.1 Gettering of metal precipitates** 557
G. Bronner
IBM T.J. Watson Res. Center, Yorktown Heights, NY, USA
J. Plummer
Stanford Univ., Stanford, CA, USA

15.00 **D2.2.2 Deep states in rapid annealed silicon** 561
M. Di Marco and A.R. Peaker
Manchester Univ., Manchester, UK
C. Hill
Plessey, Caswell, UK
M. Hart
Southampton Univ., Southampton, UK
A.E. Glaccum
British Telecom, Ipswick, UK

15.15 **D2.2.3 Gettering and deep states in p-type Czochralski silicon** 565
N. Jha and A.R. Peaker
Manchester Univ., Manchester, UK
G. Keefe-Fraundorf
Monsanto Electronic Materials, St. Louis, Missouri, USA

SESSION A2.4: CMOS TECHNOLOGY
Tuesday, September 15, 1987
SALA EUROPA - CHAIRMAN: L. BALDI

16.00 **A2.4.1 CMOS technology with self-aligned contacts and self-aligned silicide** 571
J.M. Moret, P. Weiss and H. Luginbühl
Swiss Center for Electronics and Microtechnology, Neuchâtel, Switzerland
M. Dutoit
Swiss Federal Inst. of Technology, Lausanne, Switzerland

16.15 **A2.4.2 1μm MOS devices with self-aligned titanium silicide and CVD tungsten as first metallization level** 577
C. Arena, S. Deleonibus, G. Guegan, P. Laporte, F. Martin and J.L. Pelloie
LETI/IRDI, Grenoble, France

16.30 **A2.4.3 A comparison of retrograde and conventional n-wells for sub-micron CMOS circuits** 581
A.G. Lewis, R.A. Martin, J.Y. Chen, T.Y. Huang and M.Koyanagi
Xerox, Palo Alto, CA, USA

16.45 **A2.4.4 Impact of S/D preamorphization on CMOS performance** 585
C. Mazuré, J. Winnerl and F. Neppl
Siemens, München, FRG

17.00 **A2.4.5 The influence of trench isolation on sub-micron transistors** 589
M.C. Roberts, D.J. Foster, P.H. Bolbot and P.L. Medhurst
Plessey, Caswell, UK

17.15 **A2.4.6 Electrical evaluation of SWAMI structures** 593
T. Cavioni and F. Gualandris
SGS Microelettronica, Milan, Italy

17.30 **A2.4.7 UPMOS - A new approach to submicron VLSI**
W.T. Lynch, P.D. Foo, R. Liu, J. Lebowitz, K.J. Orlowsky and S.J. Hillenius
ATT Bell Labs., Murray Hill, N.J., USA
(extended abstract not available at the time of printing)

SESSION B2.4: GaAs MESFET RELIABILITY
Tuesday, September 15, 1987
SALA ITALIA - CHAIRMAN: F. FANTINI

16.00 **B2.4.1 Temperature stability of AuGeNi ohmic contacts to GaAs** 601
A. Callegari, M. Murakami, J. Baker, Yih-Cheng Shih and D. Lacey
IBM, T.J. Watson Res. Center, Yorktown Heights, NY, USA

16.15 **B2.4.2 Thermal stability of non-alloyed ohmic contact to n-GaAs** 605
S.S. Lau and B. Zhang
Univ. of California, San Diego, La Jolla, CA, USA
A. Migliori
Bologna Univ., Dptm. of Physics, Bologna, Italy
A. Paccagnella
Trento Univ., Trento, Italy
M. Vanzi
Telettra, Bologna, Italy

16.30 **B2.4.3 GaAs-Device life-time improvements by new results on metal electrode fabrication** 609
R.P. Gupa
Central Electronics Res. Inst. Pilani, Rajasthan, India
H.L. Hartnagel, K.H. Kretschmer, R. Schütz and J. Würfl
Technische Hochschule, Darmstadt, FRG

16.45 **B2.4.4 Effects of high current and temperature in power MESFET metallizations** 613
C. Canali
Padua Univ., Italy
F. Chiussi and L. Umena
Telettra, Milan, Italy
M. Vanzi,
Telettra, Bologna, Italy
E. Zanoni
Bari Univ., Bari, Italy

17.00 **B2.4.5 Failure mechanisms study of a standard GaAs IC technology** 617
G. Kervarrec, J.M. Dumas, J.Y. Boulaire
CNET, Lannion, France
J.F. Bresse
CNET, Bagneux, France

17.15 **B2.4.6 Reliability of low-noise microwave HEMT by MOCVD** 621
K. Tanaka, J. Kobayashi, H. Takakuwa and Y. Kato
Sony Corp., Kanagawa, Japan

SALA BIANCA - CHAIRMAN: A. BROERS

16.00 **P.8** ***Ion Beam Lithography (invited)*** 625
G. Stengl, H. Löschner and E. Hammel
IMS - Ion Microfabrication Systems, Wien, Austria
E.D. Wolf
National Nanofabrication Facility, Cornell Univ., Ithaca, N.Y., USA

SESSION C2.4: POWER DEVICES II
Tuesday, September 15, 1987
SALA BIANCA - CHAIRMAN: E. RIMINI

16.30 **C2.4.1 Experiments and modeling for U.H.F. power vertical DMOS transistors** 637
G. Tardivo, A. Senes and G. Cazaubon
Thomson Semiconductors, Tours, France
P. Rossel, M. Belabadia and R. Maimouni
LAAS-CNRS, Toulouse, France

16.45 **C2.4.2 Verigrid-FCTh Switching 10A at 1000V** 641
H. Gruening and J. Voboril
BBC Brown Boveri, Baden, Switzerland

17.00 **C2.4.3 Electron energy effects in B-Irradiated power semiconductor devices** 645
P.G. Fuochi and A. Martelli
CNR-FRAE, Bologna, Italy
E. Gombia
CNR-MASPEC, Parma, Italy
C. Malfatto, B. Passerini and M. Zambelli
Ansaldo, Genoa, Italy

17.15 **C2.4.4 Helium implantation for lifetime control in silicon power devices** 649
W. Wondrak and A. Boos
AEG, Frankfurt, FRG

17.30 **C2.4.5 Isolation techniques in power integrated circuits with vertical current flow** 653
R. Zambrano, G. Ferla, S. Musumeci and M. Paparo
SGS Microelettronica, Catania, Italy

SESSION P2.2: POSTERS
16.30 **Tuesday, September 15, 1987**

a. MOS devices, measurements and special devices

P2.2.1 The small-signal behaviour of polycrystalline-silicon MOSFET's 659
A. Gnudi, P. Ciampolini, R. Guerrieri, M. Rudan and G. Baccarani
Bologna Univ., Bologna, Italy

P2.2.2 Improved determination of surface mobility at MOSFETs with thin gate oxide 663
W. Soppa and H.G. Wagemann
Technical Univ., Berlin, FRG

P2.2.3 Mobility model for silicon inversion layers 667
A.J. Walker and P.H. Woerlee
Philips Research Labs., Eindhoven. NL

P2.2.4 Leakage currents in low temperature processed polycrystalline-Si TFTs 671
S.D. Brotherton, N.D. Young and A. Gill
Philips Res. Labs., Redhill, UK

P2.2.5 Development of CMOS compatible p-channel junction-field-effect-transistors for very low noise applications 675
D. Stuch, H. Skapa, H. Vogt and G. Zimmer
Fraunhofer-Inst. for Microelectronic Circuits and Systems, Duisburg, FRG

P2.2.6 Statistical measurements of PMOS subthreshold current for 1.3 to 0.5 micron channel lenghts 679
T. Ternisien d'Ouville, R. Basset, D.T. Amm, S. Ravezzani, P. Delpech, D. Moi, M. Paoli, B. Minghetti and H. Mingam
CNET, Meylan, France

P2.2.7 Investigation of the Al-ultrathin SiO_2-Si system by comparison of theoretical and experimental current-voltage characteristics 683
B. Majkusiak, A. Jakubovski and A. Swit
Technical Univ., Warsaw, Poland

P2.2.8 Interface state analysis of MOSFET's with a modified charge-pumping technique 687
G. Przyrembel, W. Krautschneider, W. Soppa and H.G. Wagemann
Technische Univ., Berlin, FRG

P2.2.9 Monitoring the impurity profile and thickness of semiconductor layers with the channel conductance of buried channel field effect devices 691
G.J.L. Ouwerling and M. Kleefstra
Delft Univ. of Technology, Delft, NL

P2.2.10 Scanning of the entire energy gap at the Si-SiO_2 interface in MOSFETs using the conductance technique: comparison with dynamic transconductance measurements 695
H. Haddara
ENSERG, Grenoble, France
M. El-Sayed
Alexandria Univ., Egypt

P2.2.11 High resolution intrinsic MOS capacitance measurement system 699
P. Leclaire
Bull S.A., Les Clayes sous Bois, France

P2.2.12 Experimental evidence for different saturation velocities of electrons in silicon 703
B. Borchert and G. Dorda
Siemens, München, FRG

P2.2.13 Permeable base transistors with high transconductance built on LPVPE-grown silicon epilayers 707
A. Gruhle, L. Vescan, H. Beneking
Aachen Univ., Aachen, FRG

b. MOS reliability II

P2.2.14 Prediction of soft-error rate of 4 Mbit DRAM 709
W.H. Krautschneider and W. Meyberg
Siemens, München, FRG

P2.2.15 The effect of proximity of the bird's beak on aging of the thin oxide by high field current stress 713
J.C. Marchetaux, B. Doyle and A. Boudou
Bull S.A., Les Clayes sous Bois, France

P2.2.16 Simulation of stressed n-and p-channel MOSFET's: fixed oxide charges and fast interface states 717
A. Schwerin, W. Hänsch and W. Weber
Siemens, München, FRG

P2.2.17 Simulations of aging effects in MOS transistors 721
N. Bergonzoni
SGS Microelettronica, Milan, Italy
B. Doyle
Bull, Les Clayes sous Bois, France

P2.2.18 Gamma-radiation effects in CMOS transistors 725
S. Golubovic, S. Dimitrijev, D. Zupac, M. Pejovic and N. Stojadinovic
Nis Univ., Nis, Yugoslavia

P2.2.19 Temperature increase by self-heating in VLSI CMOS 729
D. Takacs and J. Trager
Siemens, München, FRG

P2.2.20 Charge limited breakdown in MOS capacitors 733
E.A. Amerasekera and D.S. Campbell
Univ. of Technology, Loughborough, UK

WEDNESDAY, SEPTEMBER 16, 1987

SALA EUROPA - CHAIRMAN: G. DECLERECK

9.00 **IP.9** ***Trends in non-volatile memory devices and technologies (invited)*** 743
H.E. Maes, J. Witters and G. Groeseneken
IMEC, Leuven, Belgium

SESSION A3.1: MOS MEMORIES
Wednesday, September 16, 1987
SALA EUROPA - CHAIRMAN: G. DECLERCK

9.45 **A3.1.1 Comparison of 11.5 μm^2 stacked capacitor and trench capacitor cells in a 1MB test memory** 757
L. Risch, W. Sesselmann and R. Tielert
Siemens, München, FRG

10.00 **A3.1.2 Isolation-related leakage in a 4Mb DRAM cell** 761
P.A. Murkin, H.M. Muehlhoff, S. Roehl, W. Meyberg, W. Mueller, W. Bergner and R. Kircher
Siemens, München, FRG

10.15 **A3.1.3 A 18 μm^2 cell for Megabit CMOS EPROM** 765
E. Camerlenghi, P. Caprara and G. Crisenza
SGS Microelettronica, Milan, Italy

10.30 **A3.1.4 A high performance p-channel EPROM cel** 769
D. Cantarelli, A. Maurelli and L. Baldi
SGS Microelettronica, Milan, Italy

SESSION B3.1: LATCH-UP
Wednesday, September 16, 1987
SALA ITALIA - CHAIRMAN: M. MONTIER

9.45 **B3.1.1 The influence of lifetime on the lateral parasitic bipolar transistors in CMOS** 775
L. Deferm, G. Romaen, C. Claeys and R. Mertens
IMEC, Leuven, Belgium

10.00 **B3.1.2 Shallow junction contacts for latch-up resistance in CMOS** 779
F. Ruddell, E. Ling, B.M. Armstrong, H.S. Gamble and S.H. Raza
The Queen's Univ., Belfast, UK

10.15 **B3.1.3 Three dimensional distribution of latch-up current in scaled CMOS structures** 783
L. Selmi, F. Venturi, E. Sangiorgi and B. Riccò
Bologna Univ., Bologna, Italy

10.30 **B3.1.4 A new SCR parameter extraction method to help design for reliability in CMOS circuits** 787
K. Erdelyi
Hungarian Academy of Sciences, Budapest, Hungary
G. Knapp
Microelectronics Co., Budapest, Hungary

SALA BIANCA - CHAIRMAN: A. WIEDER

9.45 Panel on «*Very High Speed Integrated Circuits*» 791

SESSION D3.1: LASER I
Wednesday, September 16, 1987
SALA AZZURRA - CHAIRMAN: C. WOOD

9.45 **D3.11 Degradation behaviors in LPE grown DFB laser diodes** 795
M. Fukuda, M. Suzuki, G. Motosugi and T. Ikegami
NTT Electr. Communications Labs, Kanagawa, Japan
J. Yoshida
NTT Res. and Dev. Bureau, Tokyo, Japan

10.00 **D3.1.2 1300nm DFB-laser with reactive ion beam etched grating deformation-free overgrown by LPE** 799
M. Schilling, K. Wünstel, H. Schweizer, J. Scherb, A. Mozer, K. Lösch and O. Hildebrand
Standard Elektrik Lorenz, Stuttgart, FRG

10.15 **D3.1.3 Design and fabrication of λ/4 phase jump and tunable DFB laser structures**
C.J. Armistead, B.R. Butler, S.J. Clements, A.J. Collar, D.J. Moule and S.E.A. Whiteaway
STC Technology, Harlow, UK
(extended abstract non available at the time of printing)

10.30 **D3.1.4 1.5μm DFB-BH-laser grown by hybrid LPE-MOVPE growth** 803
D. Lesterlin, J. Charil, B. Rose, M. Gilleron, P. Correc and J.C. Bouley
CNET, Bagneux, France

SALA ITALIA - CHAIRMAN: G. BENTINI

11.15 **IP.10** ***Light-guided etching for III-V semiconductor devices fabrication (invited)*** 807
D. Podlesnik
Columbia Univ., New York, USA

SALA BIANCA - CHAIRMAN: T. IKEGAMI

11.15 **IP.11** ***Design and performance of high power semiconductor lasers (invited)*** 817
K. Mettler
Siemens, München, FRG

SESSION A3.2: THIN MOS DIELECTRICS
Wednesday, September 16, 1987
SALA EUROPA - CHAIRMAN: P. BALK

11.15 **A3.2.1 Comparison of methods characterizing time dependente dielectric breakdown in thin oxide and oxide-nitride-oxide layers** 829
P. Hiergeist, M. Kerber, R. Baunach and A. Spitzer
Siemens, München, FRG

11.30 **A3.2.2 Comparison of the interfacial stress resistance in rapid thermally processed thin dielectrics** 833
R.B. Calligaro
GEC Hirst Res. Center, Wembley, UK
P.J. Rosser and P.B. Moynagh
STL Technology, Harlow, UK

11.45 **A3.2.3 Electrical characteristics and reliability of thin oxide-nitride-oxide stacked films** 837
L. Do Thanh and P. Balk
Aachen Technical Univ., Aachen, FRG

12.00 **A3.2.4 Trapping properties of very thin nitride/oxide gate insulators** 841
J.Y.C. Sun, M. Arienzo and K. Stein
IBM, T.J. Watson Res. Center, Yorktown, NY, USA
L. Dori
CNR-LAMEL, Bologna, Italy

12.15 **A3.2.5 Electron avalanche injection in thin nitrided SiO_2 films** 845
M. Severi, M. Impronta and M. Bianconi
CNR-LAMEL, Bologna, Italy

SESSION B3.2: COMPOUND SEMICONDUCTORS TECHNOLOGY III
Wednesday, September 16, 1987
SALA ITALIA - CHAIRMAN: G.F. PIACENTINI

11.45 **B3.2.1 Deposition and passivation properties of plasma CVD AlN films on GaAs using metalorganic Al source** 851
F. Hasegawa, T. Takahashi, K. Kubo, S. Ohnari, Y. Nannichi and T. Arai
Tsukuba Univ., Tsukuba Science City, Japan

12.00 **B3.2.2 An investigation of deep levels in GaAs FET's by selective de-excitation of the deep donor level EL2** 855
J. Madden, M.R. Brozel and A.R. Peaker
Manchester Univ., Inst. of Science and Technology, Manchester, UK
G. Ashcroft
Plessey, Towcester, UK

12.15 **B3.2.3 A DLTS investigation of VPE GaAs MESFET's** 859
C. Ghezzi
Parma Univ., Dipartimento di Fisica, Parma, Italy
E. Gombia, R. Mosca
CNR-MASPEC, Parma, Italy
M. Pillan
TELETTRA, Milano, Italy

SALA BIANCA - CHAIRMAN: H. WIEDER

11.45 Panel on «*Very High Speed Integrated Circuits*»

SESSION D3.2: LASER II
Wednesday, September 16, 1987

SALA AZZURRA - CHAIRMAN: J. TURNER

11.45 **D3.2.1 Analysis and design of phase-locked diode laser arrays** 865
F. Lozes-Dupuy, A. Bensoussan, S. Bonnefont, G. Vassilieff and H. Martinot
LAAS-CNRS, Toulouse, France

12.00 **D3.2.2 Reduction of losses in gain-guided laser arrays compared with single stripe lasers** 869
J. Gerner and W. Schairer
Telefunken, Heilbronn, FRG

12.15 **D3.2.3 Low threshold BH lasers emitting at 1.5 μm made from gas source MBE heterostructures** 873
B. Fernier, L. Goldstein , B. Bonnevie, D. Sigogne and J. Benoit
CGE, Laboratoires de Marcoussis, France
A. Carriere and T. Lavolee
Alcatel-CIT, Marcoussis, France

THURSDAY, SEPTEMBER 17, 1987

SALA EUROPA - CHAIRMAN: J.P. NOBLANC

9.00 **P.12** ***Integrated optics: $LiNbO_3$ or semiconductors? (invited)*** 877
M. Papuchon
Thomson-CSF, Orsay, France

SESSION A4.1: INTEGRATED OPTOELECTRONICS I
Thursday, September 17, 1987
SALA EUROPA - CHAIRMAN: M. ROCCHI

9.45 **A4.1.1 1.3 μm buried ridge laser grown on S.I.INP substrate: a structure suitable for high speed modulation and optoelectronic integration**
P. Devoldere, A. Paraskevopoulos and A. Godin
CNET, Bagneux, France
(extended abstract not available at the time of printing)

10.00 **A4.1.2 Current-injection analysis of invertible InGaAsP/InP double hetero-structure bipolar transistors** 883
H.G. Bach, F. Fiedler and N. Grote
Heinrich-Hertz-Inst. für Nachrichtentechnik, Berlin, FRG

10.15 **A4.1.3 Heterojunction bipolar phototransistors for high speed logic and communication applications** 887
A.J. Doherty and W.S. Truscott
Manchester Univ., Inst. of Science and Technology, Manchester, UK

10.30 **A4.1.4 AlInAs/GaInAs MESFET grown by MBE** 891
L. Giraudet, M. Allovon, J. Ch. Renaud and A. Scavennec
CNET, Bagneux, France

SESSION B4.1: BIPOLAR MODELLING I
Thursday, September 17, 1987
SALA ITALIA - CHAIRMAN: H.C. DE GRAAFF

9.45 **B4.1.1 The influence of emitter sidewall injection on transistor noise figure** 897
H.R. Claessen, J.A.M. Geelen and H.C. de Graaff
Philips Research Labs., Eindhoven, NL

10.00 **B4.1.2 Carrier multiplication and avalanche breakdown in self-aligned bipolar transistors** 901
M. Reisch
Siemens, München, FRG

10.15 **B4.1.3 Electrical characterization of polysilicon/monosilicon interface** 905
S. Bellone and P. Spirito
Naples Univ., Dptm. of Electronics Engineering, Naples, Italy
M. Arienzo
IBM T.J. Watson Res. Center, Yorktown Heights, NY, USA

10.30 **B4.1.4 Current gain dependence on the emitter size of polysilicon-emitter bipolar transistors** 909
M. Miura-Mattausch
Siemens, München, FRG

SESSION C4.1: TEST CHIPS
Thursday, September 17, 1987
SALA BIANCA - CHAIRMAN: J. ROBERTSON

9.45 **C4.1.1 The effect of device geometry on IGFET characteristics** 915
J.A. Serack, A.J. Walton and J.M. Robertson
Edimburgh Univ., Edimburgh, UK

10.00 **C4.1.2 The measurement of transistor characteristics using on-chip switching for the connection of instrumentation** 919
D. Ward, A.J. Walton and J.M. Robertson
Edimburgh Univ., Edimburgh, UK

10.15 **C4.1.3 A 2D carrier profiling technique for VLSI planar structures** 923
C. Hill, P.J. Pearson, B. Lewis, A.J. Holden and R.W. Allen
Plessey, Caswell, UK

10.30 **C4.1.4 Statistical defectivity control for VLSI devices** 927
R. Traversini, A. De Lisio, M. Tosi and G. Barbuscia
SGS Microelettronica, Milan, Italy

SALA ITALIA - CHAIRMAN: H. RYSSEL

11.15 **IP.13** ***The physics of silicide base transistors (invited)*** 931
E. Rosencher, F. Arnaud d'Avitaya, P.A. Badoz, G. Glastre and G. Vincent
CNET, Maylan, France

SESSION A4.2: INTEGRATED OPTOELECTRONICS II
Thursday, September 17, 1987
SALA EUROPA - CHAIRMAN: V. GHERGIA

11.15 **A4.2.1 Process and model of short-gate diffused InGaAs JFET's for integrated PIN-FET photodetector** 943
L. Nguyen, M. Allovon, P. Blanconnier, E. Caquot and A. Scavennec
CNET, Bagneux, France
B. Bourdon
Laboratoires de Marcoussis CRCGE, France

11.30 **A4.2.2 Monolithic integration of a GaInAs JFET and a GaInAs photodiode** 947
P.J.G. Dawe, D.A.H. Spear, G.H.B. Thompson and G.R. Antell
STC Technology, Harlow, UK

11.45 **A4.2.3 InGaAs/InP SAM photodetectors fabricated by structure controlled acceptor diffusion** 951
E. Kühn, N. Klötzer, H.W. Marten and A. Schmiech
Stuttgart Univ., Stuttgart, FRG

12.00 **A4.2.4 Monolithic integration of a GaInAs photoconductor with a n^-/n^+ GaAs RIB waveguide: a simple design device** 955
F. Mallecot, J.P. Vilcot and D. Decoster
Lille Univ., Villeneuve d'Ascq, France
M. Razeghi
Thompson-CSF, Orsay, France

12.15 **A4.2.5 MBE growth and processing of InGaAs/InGaAlAs/InP monolithical integrated ridge waveguide photodiodes** 959
P. Cinguino, F. Genova, C. Rigo, C. Cacciatore and A. Stano
CSELT, Turin, Italy

SESSION B4.2: BIPOLAR MODELLING II
Thursday, September 17, 1987
SALA ITALIA - CHAIRMAN: H.C. DE GRAAFF

11.45 **B4.2.1 A new simple analytical evaluation of minority carrier current in arbitrarily doped region with non-thermal generation of carriers** 965
C.R. Selvakumar and D.J. Roulston
Waterloo Univ., Dptm. of Electrical Engineering, Waterloo, Canada

12.00 **B4.2.2 Measurement and calculation of base resistance components of modern high speed bipolar transistors** 969
J. Fertsch, H. Voit, H. Klose and W.R. Böhm
Siemens, München, FRG

12.15 **B4.2.3 A bipolar DC model for transistors fabricated in a CMOS process** 973
D.J.F. Doyle and W. Lane
University College, Cork, Ireland

SESSION P4.1: POSTERS
11.45 **Thursday, September 17, 1987**

a. GaAs technology and devices

P4.1.1 RBS and XPS depth profiling of rapid thermally alloyed AuGe/Ni Contacts to GaAs
A.G. Nassibian, D.V. Morgan, M.S. Frost and C.E.C. Wood
GEC, Wembley, UK
(extended abstract not available at the time of printing)

P4.1.2 Emitter concentration and interface effects in the base-emitter characteristics of AlGaAs/GaAs heterojunction bipolar transistors 979
P. Camara, E. Muñoz, I. Izpura, J. Lablanca and E. Lapeña
ETSI Telecomunication, Univ. Politecnica Madrid, Spain
M.A. Pate, G. Hill, P. Mistry, J.S. Roberts and H.Y. Hall
Sheffield Univ., Dptm. of Electronic Engineering, Sheffield, UK

P4.1.3 A computer aided interpretation of mobility profile measurements in gallium arsenide FET structures 983
P. Pillan and F. Vidimari
Telettra, Milan, Italy

P4.1.4 Planar GaAs mixer diodes for millimeter wave MMICs 987
J. Selders and A. Colquhoun
Telefunken, Heilbronn, FRG
K.E. Schmegner
AEG, Ulm, FRG

P4.1.5 Low-noise bulk unipolar devices in Si and GaAs 991
H. Beneking, J.M. Cloos, G. Fernholz, M. Marso, P. Roentgen and L. Vescan
Aachen Technical Univ., Aachen, FRG

P4.1.6 Tungsten silicide resistors for GaAs MMICs 995
D.A. Allan, T.K. Ng and M.J. Gilbert
British Telecom, Ipswich, UK

P4.1.7 MESFETs on n-GaInAs with barrier enhanced Schottky gates 999
G. Fernholz, W. Lange, R. Westphalen, P. Balk and H. Beneking
Aachen technical Univ., Aachen, FRG

b. Optoelectronics

P4.1.8 Conduction mechanisms analysis and simulation of 1.3 μm laser diodes 1003
S. Mottet and J.E. Viallet
CNET, Lannion, France
A. Changenet
CNET, Paris, France
E. Dudda and A. Accard
CGE-LdM, Marcoussis, France
R. Blondeau and M. Krakowski
Thomson CSF, Orsay, France

P4.1.9 Failure analysis of GaAlAs emitters 1007
G. Conte and F. Magistrali
Telettra, Milan, Italy
F. Fantini and M. Vanzi
Telettra, Bologna, Italy

P4.1.10 Technological critical points of InGaAsP/InP 1.3 μm lasers as evidence of 12000 hours CW operating life tests 1011
R. De Franceschi, M. Liberatore, P. Montangero and A. Piccirillo
CSELT, Torino, Italy

P4.1.11 Turn-on delay time fluctuations in unbiased gain-switched AlGaAs-GaAs multiple quantum well lasers 1017
E.H. Böttcher, K. Ketterer and D. Bimberg
Berlin Technical Univ., Berlin, FRG

P4.1.12 Carrier density and inter-valance band absorption in InGaAs(P) lasers 1021
S. Hausser, E. Zielinski, M. Asada and H. Schweizer
Stuttgart Univ., Stuttgart, FRG
H. Burkhard and E. Kuphal
Forschungsinstitut der Deutschen Bundespost, Darmstadt, FRG

P4.1.13 Two-dimensional model for cleaved-coupled cavity and external cavity lasers 1025
J.P. Van de Capelle, R. Baets and P.E. Lagasse
IMEC, University of Ghent, Belgium

P4.1.14 Simple model for semiconductor laser end facet reflectivity 1029
G.P. Bava and A. Bianco
Politecnico, Torino, Italy
I. Montrosset
Genoa Univ., Genoa, Italy

P4.1.15 A tunable twin-stripe distributed feedback laser model 1033
M. Federighi
Marconi Italiana, Genoa, Italy
T. Kumar
GEC, Wembley, UK

SALA EUROPA - CHAIRMAN: B.L.M. WILSON

14.00 **IP.14** ***AlGaAs/GaAs modulation doped FETS for ultra-high speed signal processing applications (invited)*** 1037
A. Christou
Crete Univ., Iraklio, Greece

SESSION A4.3: BIPOLAR TECHNOLOGY
Thursday, September 17, 1987
SALA EUROPA - CHAIRMAN: A. WIEDER

14.45 **A4.3.1 Double self-aligned bipolar transistors using salicide contacts** 1047
H. Kabza, M. Reisch, V. Probst, W. Böhm, J. Fertsch, H. Schaber and H. Eggers
Siemens, München, FRG

15.00 **A4.3.2 Speed-power relation of modern bipolar technology** 1051
P. Weger
Siemens, München, FRG
H.M. Rein
Ruhr University, Bochum, FRG

15.15 **A4.3.3 Study on current transport mechanism in amorphous SiC emitter HBT** 1055
K. Sasaki and S. Furukawa
Tokyo Inst. of Technology, Yokohama, Japan

15.30 **A4.3.4 Device design of a high voltage BiCMOS IC** 1059
Tong Qin-Yi and Wu Wei
Nanjing Inst. of Technology, Nanjing, China

SESSION B4.3: ULTRA FAST OPTOELECTRONICS
Thursday, September 17, 1987
SALA ITALIA - CHAIRMAN: H. MELCHIOR

14.45 **B4.3.1 Laser diode with an integrated gain/loss modulator for the generation of picosecond optical pulses by active mode locking** 1065
J. Werner, G. Guekos and H. Melchior
Swiss Federal Inst. of Technology, Zurich, Switzerland

15.00 **B4.3.2 Indium tin oxide-gallium arsenide photodiodes for operation at frequencies beyond 110 GHz** 1069
D.G. Parker
GEC Research, Wembley, Middlesex, UK
W. Sibbett
Dept. Physics, Univ. St. Andrews, UK

15.15 **B4.3.3 High speed control of microwave signals using InP:Fe photoconductive devices** 1073
I. Andersson and S.T. Eng
Chalmers Univ., Gothenburg, Sweden

15.30 **B4.3.4 Epitaxial silicon avalance photodiodes for single photon detection with picosecond resolution** 1077
S. Cova, G. Ripamonti, A. Lacaita and M. Ghioni
Politecnico of Milan, Italy

SESSION C4.2: LATE NEWS
Thursday, September 17, 1987
SALA BIANCA - CHAIRMAN: P.U. CALZOLARI

IP.1. - INVITED PAPER

BI(C)MOS

DREAM OR NIGHTMARE?

P.A.H. HART

Philips Research Laboratories
P.O. BOX 80.000
5600 JA Eindhoven, The Netherlands

The paper presents a summary of advantages of mixed bipolar-CMOS processes for analog and digital VLSI and some design considerations.

INTRODUCTION

In recent years a strong interest in BI(C)MOS has grown. BI(C)MOS is a technology combining bipolar and (C)MOS on a single chip. Because it now is possible to exploit the strenghts of these disciplines in a single integrated circuit, it is in a way a dream come true for electronics designers. In the past, however, the penalty, that is a more complex process, has restricted such technologies to rather specialized applications. In modern technologies both bipolar and MOS have grown to become very similar and both have become rather complex. Therefore, now that it proves possible to use a large number of steps in common such a process can be economically viable and need no longer be a technologists nightmare. In fact, estimates are in the order of a 20% increase of wafer price which will be offset by a considerable enhancement of performance, increase in yield and occasionally a 35-45% reduction in circuit complexity (1) and a decrease of chip area by more than a factor of 2, cf (2).

In this paper we shall present a summary of advantages of mixed processes, a survey of the general status and consider technologies and device structures. However, before one can start doing this, one has to define what will will be covered by the name BI(C)MOS as many authors use the name for different kinds of processes and devices or application areas. BIMOS or BIDFET appears to be used in the early days in the context of mixed processes such as used for circuits like operational amplifiers. Later on it was used in the context of special power processes and smart power devices. In this paper we shall use it in the context of general purpose processes for VLSI with some side observations on power. An important aspect here is to exploit the performance advantages of bipolar and CMOS and at the same time use the higher yield capability of MOS to reach VLSI levels more easily.

DEVICE ASPECTS

Table I lists some important properties of bipolar and CMOS transistors in circuits

From table I it can be deduced that generally for devices bipolar is in an advantageous position for analog applications because of the better gain and low wideband noise. CMOS obviously is more attractive for digital control and data processing functions because of its low quiescent power, good speed and generally good packing density. The mixture of bipolar and MOS offers unique advantages, however, in both the analog and the digital

Table I

Strengths of Bipolar and CMOS.

Bipolar	(C)MOS
large transconductance	high impedance
linear h_{FE}	zero dc gate current
exponential characteristic	near quadratic
low voltage offset	zero dc dissipation
low supply voltage	low narrow band noise
low 1/f noise	high slew rate
low wide-band noise	no second breakdown
fast	self isolating
low logic swing	no avalanche breakdown
good capacitor drive capability	reduction
no hot electron limit	

fields. It features for instance, high-impedance zero dc gate current high gain operational opamps, switch capacitor filters and gyrators. The latter using MOS for large time constants. Furthermore, precision pairs in mixed AD and DA systems in the field of analog applications are possible. In digital applications mixed sense-amplifiers (3) and buffers (4,5) can significantly increase the capability of CMOS with regard to speed and compactness. Speed improvements up to a factor of two with respect to their CMOS counterparts are reported, cf. (4), fig. 1.
In the context of electronic advantages it should be noted that some problems can also occur. Because bipolar transistors can inject into the substrate one has to take measures to prevent latch-up. Furthermore CMOS logic generates considerable transient noise which must be prevented from entering into an adjacent sensitive analog part. Both problems can effectively be suppressed by proper measures.

Returning briefly to table I then one observes that in the absence of stored charges

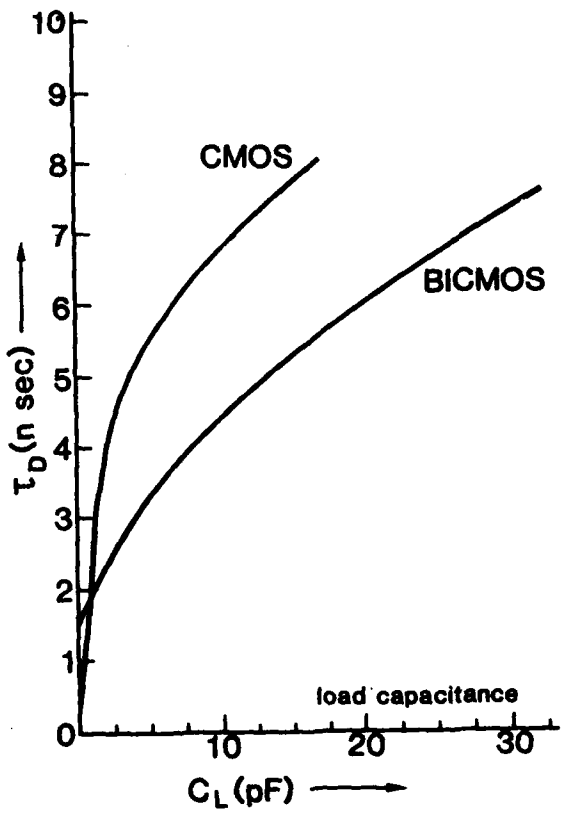


Fig. 1. Comparison delay time of BICMOS buffers and CMOS buffers as a function of load capacitance C_L.

MOS offers a better slew rate. This is advantageous in operational amplifiers. Moreover, as there is no avalanche-induced breakdown reduction in MOS this makes the device attractive for high-voltage application. Such a fast and high voltage amplifier is described in ref (6). Particularly if one uses DMOS this leads to mixed processes suitable for high voltage (200V and above) and for power cf (7).

THE PROCESS

There are several ways to arrive at a combined process, a very common one being to start from a CMOS process. One could of course also start from a purely bipolar process or even from scratch. Here we shall follow the first line of thought we start from N-well CMOS. The left of fig. 2a depicts a cross section of CMOS. The customary source and drain diffusions seen are typically in the order of a few tenths of a micron deep, those of the N-well in the order of two or three microns. If one wants to incorporate an NPN bipolar transistor one first has to provide an additional P-diffusion for the base cf. fig. 2a. This transistor, however, exhibits a high collector series resistance. Furthermore, because of the overall structure being surrounded by lowly doped regions it will be rather prone to latch-up . Such a "simple" transistor is difficult to design with. A much better result is obtained with structures shown in fig. 2b. Here with respect to that of fig. 2a an N^+ buried layer touching the N-well and an epitaxial P-layer is used. This adds two masking steps on top of the CMOS process having typically around 10-12 masking steps for a modern (single metal) process. In principle in this process the NPN transistor, being surrounded by P material, is isolated. Latch-up via the bottom of the NPN is largely prevented by the N^+ buried layer, which is also underneath the PMOS. Undesirable lateral action is still possible however, particularly as the P epilayer will be rather high-ohmic. This can, at the cost of extra steps, of course be prevented by a normal deep P^+ diffusion, deep oxide isolation or trenches. A very simple solution, however, is shown in fig. 2c. It is possible by self alignment, i.e. no extra masking, to have a P^+ buried layer adjacent to the N^+ layer and to use the P^+ diffusion of the CMOS process on top to supress possible channels. This way a compact and largely latch-up free structure is obtained. In fig. 3 a profile of the NPN transistor (8) and in fig. 4 a cross section is shown of an isolation structure as is used by Philips in a 2 µm lithography BIC)MOS process (2). In table II the effect of a buried layer obtained by Siemens (9) but in a 1.4 µm process regarding collector series resistance and a substrate paras itic

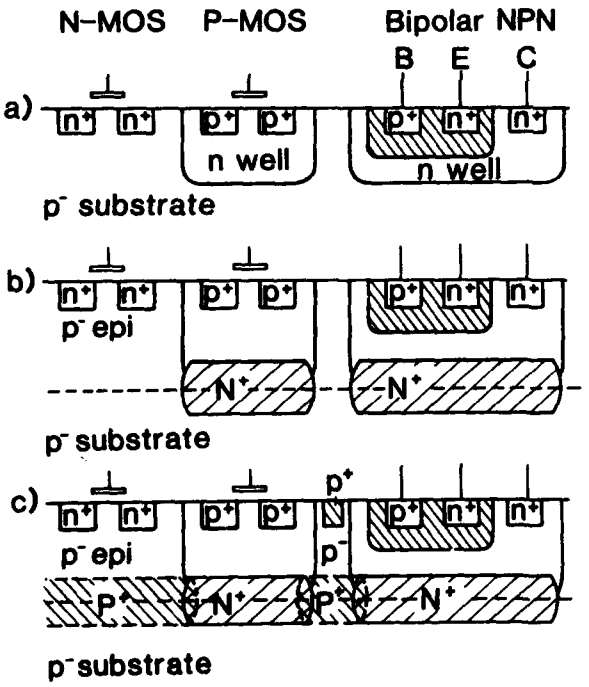


Fig. 2. BICMOS development from N-well CMOS.

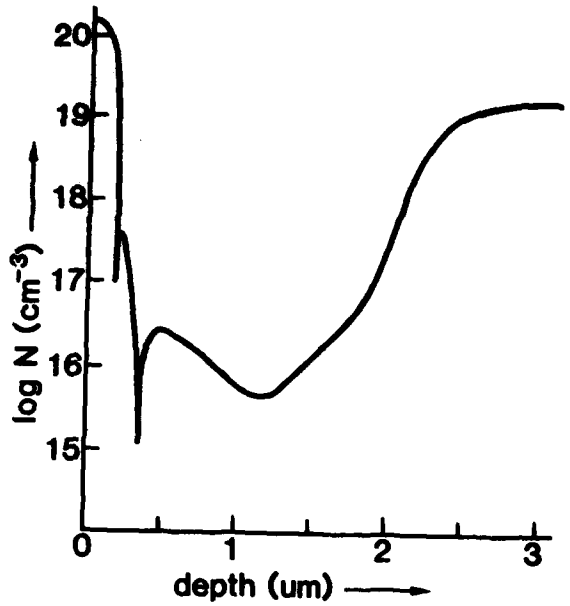


Fig. 3. Impurity profile of NPN, 1.5 µm process. Note : upwards slope in the collector. Typically achieved by also using phosphorus in the N^+ buried layer.

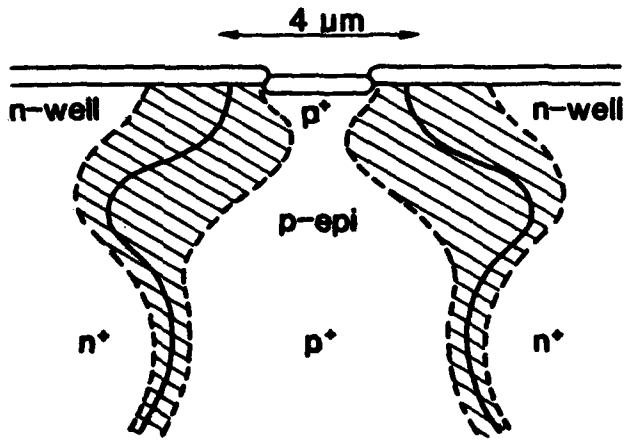


Fig. 4. Isolation structure, hatched areas are depletion layers at 18V. Without the p^+ buried layer the lateral distance should be 9.5 µm instead of 4.5 µm. Similarly omission of the p^+ channelstop diffusion requires the distance to be at least 6.5 µm. Epi width$\sim$ 4 µm.

Table II

Effect of the buried layer (9)

	"No" layer fig. 2a	"With" layer fig. 2b
H_{FEmax}	60	80
R_C (Ω)	200	30
H_{FE} pnp	25	1

Emitter size 4x30 µm.

(latch-up) pnp are presented for different structures. Latch-up is studied by many authors cf (11).

The BI(C)MOS technology is being developed at many places and significant headway has been made. An overview is given in table III. This overview does not claim to be complete it is only intended to present a trend, namely that most manufacturers use structures similar to fig. 2b or fig. 2c, sometimes with more refinements, rather than the simplest but more difficult to design fig. 2a option. Many use N-well CMOS, but several use P-well or start from bipolar rather than from CMOS. Standard "VLSI" BI(C)MOS needs 14-16 masks, options may increase this to 20.

Table III

BICMOS

Company / Reference	Digital	Analog	*) Type	Litho (µm)	N° Masks	F_T NPN (GHz)	Remarks
DEC 5	+	-	2	1.5	12		
Honeywell 5	+	+	1	3, 1.25	14		BICMOS II, III
Hughes 5	+	+	3	1.5	20	5	pnp 2.5 GHz
Motorola 5	+	+	3	1.75	15		Dyn PLA's
National 5,10	+	+	2	2,3		1.9	W fuses
Texas Instr. 18	-	+	3		12		BIDFET, DMOS 280V
RCA 15	+	+	1	3		0.7	QMOS, Power 60V
Philips 7	-	+	3	6	12		DMOS 200V
Philips 2	-	+	3	2	14	3.5	
12	+	+	3	1.5	16	5	
Siemens 12	+	-	2	1.4	14	2.6	
Telefunken 5	+	+	3	2	14	1	pnp
	+	+	3	1	16	2	pnp
Thomson 16	+	+	3	4	13	0.4	120V HV MOS
SGS-ATES 5,4		+	3		13	1	DMOS HV 60V
Hitachi 5	+	+	3	3		3	ABC, pnp + I^2L
" 5,4	+	+	3	2		4	Hi-BICMOS
" 5,13	+	+	3	1.3		9	poly Si emitter 0.4 nS/gate
" 5	-	+	3	5		0.4	20V
NEC 5	+		3	1.6		6	poly Si emitter
RICOH 5	+	+	1		10		40V PMOS + 60V NPN metal gate.
Toshiba 3,14	+		1,2	1.2,1			MoSi gate, option N^+ buried layer

*) Type 1. No epi.
2. Epi + buried layer.
3. More sophisticated.

PERFORMANCE

We shall now return to device properties and their influence on circuit performance. If one follows the approach as sketched above then one starts from a standard N-well CMOS process. This offers many advantages, as usually there is not only the process itself available, but also the associated culture of CMOS CAD and electronic design. However, such an approach harbours a constraint in terms of available diffusions and last but not least in dope and depth of the N-well. A typical profile of a NPN is shown in fig. 3. Apart from the diffusions, the properties of the bipolar transistor are determined by the width and dope of the epilayer. In very high performance bipolar the epilayer is chosen as thin as possible, say 1 μm or even less. This is thinner than the usual N-well by a factor of two. If one brings the N^+ layer which forms the bottom of the N-well to close to the source and drain areas then this degrades the transistor by increasing capacitance and the body effects, fig. 5. To avoid this a sufficiently wide epilayer is needed. This has important consequences for the bipolar part; maximum f_T will be lower than in a high-speed "bipolar only" process. At the same time however, a better Early effect and better lateral pnp's are possible. These are essential features for analog. In fig. 6 such a trade off as obtained by Rausch et al (2) is illustrated.

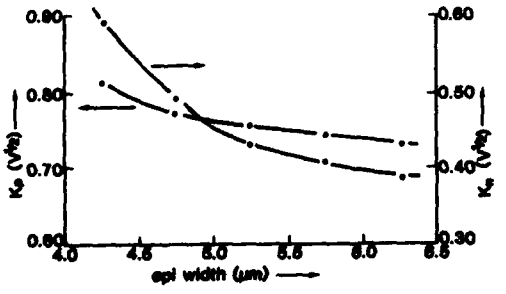


Fig. 5. Body effect of pMOS and nMOS measured as a function of epiwidth.

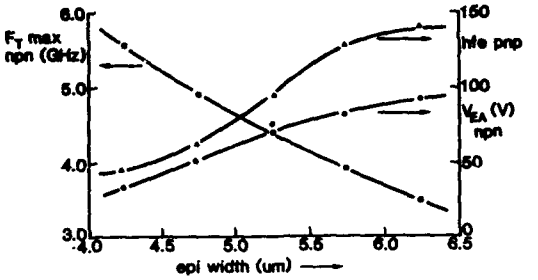


Fig. 6. Measured dependence of bipolar device parameters as a function of epi width.

FURTHER DEVELOPMENTS

We shall now try to broaden the picture to include further developments. A very conspicuous element is of course lithographic shrink which in the future will continue down to 0.5 μm and even less. Such a shrink has several consequences in a mixed process. Starting again from a given MOS process shrink means a reduction of supply voltage, although this can be somewhat postponed by the use of graded drains or use of LDD structures. Another strong effect is the increase of dope in the substrate or in the pocket to prevent punch-through.

A reduction of supply voltage will affect the systems in which the circuit is to be used in much the same way as is the case for plain CMOS. With regard to a possible analog part this means a reduction of dynamic range which occasionally might be unacceptable. Assuming for the present that the supply voltage is acceptable one still has to discuss the consequences of the dope level increase. For the NPN transistor as well a for MOS the capacitances per area will increase as the depletion regions are depressed. At the same time it becomes possible to decrease the epilayer width with no additional sacrifice in body factor. A higher value of f_T is thereby achieved. This trend can be seen in table IV.

Such an NPN transistor is not always suitable for high-quality analog as the

Table IV

Effect of lithographic scaling and N-well

Litho	2 μm	1.5 μm	1.5 μm lowly doped N-well
epi width	5	2	2
K_n ($V^{1/2}$)	0.75	0.5	-
K_p ($V^{1/2}$)	0.5	0.65	-
$R_C C_{csub}$ (pS)		49	164
$R_B C_{CB}$ (pS)	73	37	14
F_T (GHz)	3	6	4
$H_{FE}.V_{EA}$ (V)	10000	2400	9000

product of $h_{FE}V_{EA}$ tends to be too low. The remedy is to use two different N-wells. The NPN transistor sits in a more lightly doped N-well. In the third column of table IV the relevant quantities are shown. In addition to a considerable improvement (with respect of the second column), in Early effect, also the $R_B C_{CB}$ product is improved at the cost of f_T. This makes the transistor much more suitable for analog.

For digital the picture is different $h_{FE}V_{EA}$ does not play a very significant role and the single N-well concept is therefore suitable for digital. Even more so because of the higher dope levels the f_T value of the transistor tends to increase. This will help to move the crossover point in fig. 1 to the left and the speed of the circuit will come closer to the intrinsic speed of the "nearly" unloaded Mos.
It should be remarked that apart from the lithographic shrink there will be several other improvements. One of these is quite likely the use of poly silicon in the bipolar part cf. Table III. Another could be the introduction of new isolation schemes and silicides.

CONCLUSION

In this paper we have discussed the combination of bipolar and MOS. This combination has great advantages for digital and also allows development of high-quality analog. The developments are economically viable and therefore more of a pleasant dream come true than a nightmare.

It should finally be remarked that some of the results quoted are obtained in a Philips-Siemens ESPRIT project.

REFERENCES

1. I. Fukushima, K. Kuwahara, K. Hoya, N. Horie, K. Itoigawa, S. Ichimura, M. Nagata. A BIMOS FET Processor for VCR audio, IEEE, IESCC 1983, p. 242, 243.

2. F. Rausch, H. Lindeman, W.J.M.J. Josquin, D. de Lang and P.J.W. Jochems. An Analog Bimos technology. Extended abstracts of the 18th Conference on Solid State Devices and Materials Tokyo 1986, pp. 65-68.

3. J. Miyamoto, S. Saitoh, H. Momose, H. Shibata, K. Kanzaki, T. Iizuka, A 28 nS CMOS SRAM with bipolar sense amplifiers IEEE, ISSCC 84, p. 224-225 and 344.

4. T. Hotta, I. Masuda, H. Maejima, A. Hotta, CMOS/Bipolar circuits for 60 MHz digital processing, IEEE, ISSCC 1986, p. 190-191.

5. B.C. Cole, Mixed-process chips are about to hit the big time. Electronics, March 3, 1986, p. 27-31.

6. P.G. Blanken and P. van der Zee IEEE Trans CE-31, 1985, p. 109

7. A.W. Ludikhuize, A versatile 250/300V process for analog and switching applications. IEEE Trans ED, Vol. 33, 1986, p. 2008-2015.

8. D. de Lang and W.J.M.J. Josquin, Optimization of a 1.5 μm BICMOS process. BICMOS Symposion, Electrochem Soc. Philadelphia, 1987.

9. J. Winnerl and E.P. Jacobs, Essderc 1985 and private communication BICMOS ESPRIT project.

10. R.J. Murphy, T. McFarlane, C. Sporck, K. Rapp, R. Smolen, W. Collins, R. Ramus, M. Millholan, J. Readdic. A bipolar-CMOS field programmable Array, IEEE, ISSCC 1986, p. 70-71.

11. J.A. Seitchik, A. Chatterjee and P. Yang. An analytical model for latch-up in epitaxial CMOS. IEEE ED Letters Vol EDL no. 4 April 1987, 157-159. (holding voltage PIN).

12. Part of a Philips-Siemens Esprit project.

13. A. Watanabe, T. Ikeda, T. Nagano, N. Momma, Y. Nishio, N. Tamba, M. Odaka and K. Ogine.
High speed BICMOS VLSI technology with buried twin well structure. IEEE, IEDM 1985 Washington DC, p. 423-426.

14. J. Miyamoto, S. Sautoh, H. Momose, H. Shibata, K. Kanzaki and S. Kohyama.
A 1.0 μm N-well CMOS/Bipolar technology for VLSI circuits, IEEE, IEDM 1983 Washington, p. 63-66.

15. G.M. Dolny, O.H. Schade, B. Goldsmith and L.A. Goodman,
Enhanced CMOS for analog-digital power IC applications IEEE, Trans-ED, Vol 33, 1986, p. 1985-1991.

16. G. Thomas, G. Trousset and F. Vialettes, High voltage technology offers new solutions for interface integrated circuits. IEEE Trans-ED, Vol 33, 1986, p. 2016-2024.

17. A. Andreini, C. Contiero and P. Gabiati, A New integrated silicon gate technology combining bipolar linear, CMOS logic and DMOS power parts. IEEE Trans ED, Vol 33, 1986, p. 2025-2030.

18. BIDFET, marriage of three technologies. Electronics July 7, 1983, p. 107.

NEW DEVELOPEMENTS IN SOLID-STATE DETECTORS

Emilio GATTI, Antonio LONGONI and Marco SAMPIETRO

Politecnico di Milano, Dipartimento di Elettronica and
Centro di Elettronica Quantistica e Strumentazione Elettronica CNR
Piazza Leonardo da Vinci 32, Milano 20133, Italy

The interest in semiconductor radiation detectors for energy and position measurements is rapidly increasing in the fields of high energy physics, astronomy, space applications and medicine. After a brief review of the field, the recent developements are here presented.

1. INTRODUCTION

Among the detectors which have been used in the field of nuclear and elementary particles physics, semiconductors detectors, germanium or silicon, played an important role starting from 1951,when Kenneth McKay of the Bell Telephone laboratories showed the first alpha particles pulses[1]. These pulses were due to electrons and holes generated in the depletion region of a reverse biased germanium p-n junction, seprated by the depleting field and collected at the junction contacts.

This principle has been used in every semiconductor detector since then. In fact, except for some recent developements described later in this paper, semiconductor detectors are essentially silicon or germanium p-n or p-i-n diodes, reverse biased, generally fully depleted, of different shapes and dimensions.

The fortune of semiconductor detectors is their intrinsic energy resolution due to the amount of charge released by impinging particles. For comparison, an ionizing particle in a plastic scintillator must spend 300 eV per photoelectron emitted at the cathode of an electron multiplier; a gas detector needs 30 eV spent to generate an electron-ion pair,while in semiconductors an electron-hole pair is generated with only 2.98 eV in germanium at 77 K and 3.6 eV in silicon at room temperature.

Further, for a minimum ionizing particle, due to the short range of electrons and photons in silicon,electrons and holes are generated within a column of diameter less than 1 micrometer along the trajectory with a linear density of 85 couples/micrometer.

Ge detectors are particularly suitable for X and γ spectroscopy because of the higher atomic number (Z=32) with respect to Si (Z=14) which leads to an efficiency for the photoelectric effect higher by a factor of $(32/14)^4=27.3$. However,the higher leakage current of Ge detectors with respect to silicon,due to the lower energy gap, requires operation at liquid nitrogen temperature.

Silicon,when possible, is used because of the higly developped planar technology which has no parallel for germanium. Semiconductor detector technology makes generally use of low doped materials (10^{11}-10^{13} cm^{-3}) in order to get large depletion layers (10μm-5cm) with low voltages,leading to fields below the breakdown threshold. Further requirement is to obtain in the final device,after all technological steps, a long lifetime for minority carriers (several ms) which ensures low leakage current at the collector,and a low concentration of deep trapping centers which could trap a fraction of the carriers travelling to the collector and release them with a sensible delay.

New developements in semiconductor detectors have been recently stimulated by the following two requirement from the high energy physics community:

(a) the need of an higher signal to noise ratio, in order to improve energy and timing resolution, with large area detectors and in conditions of high rates of events to be detected

(b) the need to get from a detector on a single wafer high precision measurements of the position of incidence of the ionizing particles.

The energy resolution is usually expressed in "equivalent noise charge" (ENC), that is the charge which, delivered as a delta pulse by the detector, cause at the output of the amplifying circuits connected to the detector (preamplifier plus filtering amplifier) a pulse equal to the RMS noise.

The detector parameters which determine ENC are the leakage current at the anode I_l, as source of the parallel noise current, and the anode capacitance C_d. The parameter of the electronic circuit which determine ENC are the series noise of the input FET of the preamplifier, its input capacitance C_i and its badwidth f_t.

It can be shown that, for every detector and preamplifier configuration, it is possible to find a linear filter which minimize the ENC (optimum resolution ENC_{opt}) [2,3]. The optimum resolution can be expressed as:

(1) $$ENC_{opt} \doteq (I_l C_d / f_t)^{1/4} \cdot ((C_d/C_i)^{1/2} + (C_i/C_d)^{1/2})^{1/2}$$

The ENC_{opt} is obtained with an optimum filtering which produces a pulse, at the output of the preamplifier-filtering amplifier chain, having a time length τ_{opt} :

(2) $$\tau_{opt} \doteq (C_d/(f_t I_l))^{1/2} \cdot ((C_d/C_i)^{1/2} + (C_i/C_d)^{1/2})$$

From (1) and (2) it is clear that both ENC_{opt} and τ_{opt} are minimized when input capacitance is matched to detector capacitance, i.e. $C_d = C_i$.

As the input capacitance of a FET is of the order of less than a pF, while the capacitance of a conventional p-n detector is generally greater than tens of pF (for instance the capacitance of 1 cm^2 silicon detector 300 μm thick completely depleted is of the order of 30pF) it is evident from (1) and (2) that is useful an effort in the direction of the reduction of the detector capacitance (without sacrifying active volume of the detector) in order to get a small ENC and a small resolving time (which is useful when high rates of events have to be processed).

The basic idea, in order to acheive such a result, is to conceive a detector having a collecting electrode whose dimensions are independent and much smaller than the active area of the detector. In this case the output capacitance of the detector is much smaller than that of a conventional p-n detector of equivalent active area, in which the dimensions of the collecting electrode are practically coincident with the active area.

It is possible to get such result by conceiving a device in which the motion of the charges toward the collecting electrode is practically unaffected by its geometry and voltage. The carriers are brougth to this electrode by electrical fields (static or dynamic), parallel to the surface of the detector and independently superimposed to the depleting field. Furthermore, the knowledge of the motion of carriers toward the collecting electrode, can give an information about the position of interaction of the ionizing particle.

Proceeding in this way, it is possible to design detectors whose output capacitance is comparable or even lesser than that of a good input FET of the preamplifier (detector have been built with output capacitance less than 100fF). In this conditions the limiting factor to a further improvement of resolution and resolving time are the stray capacitances of the connections between detector and preamplifier. It becomes therefore imperative the integration of the preamplifier itself directely on the detector.

Two different kind of devices are based on

these principles. One is the Semiconductor Drift Chamber (SDC), a completely new device in which the electrons are drifted to the anode by static fields. The other is the Charge Coupled Device (CCD), already in use for detection in the visible region, recentely applied in high energy phisics as position sensitive detector.

2. SILICON AND GERMANIUM TELESCOPES

Figure 1 shows a typical silicon detector implemented as a simple p-n junction with implanted p^+ and n^+ contacts on an n-type bulk. Electrons and holes, generated by the incident particle, drifting in the depletion field induce at the electrodes a current pulse. The charge preamplifier integrates the current and gives at the output a voltage proportional to the total charge collected. Collection times are typically of the order of 10 ns.

Figure 2 shows an elaborate experiment done at CERN with several detectors of this kind, mounted one behind the other, forming a so called "telescope", preceded by a particular germanium detector[5]. Each silicon detector of the "telescope" detects the charges generated by the particles traversing it. The histogram of these charges (Figure 2b) allows to determine the molteplicity of the interactions and where they occurred. It is interesting to remark the double function of the silicon as a nuclear target and as a detector, so that these detectors are called "active targets".

The germanium bulk detector, like the silicon telescope, is an active target. The device can be thought of as a sequence of equal detectors with no physical separation in between. The beam incident is parallel to the strip plane. Therefore the transversal dimension must be large enough to cover the beam section. The detector thickness is .5 cm, its width .5 cm and its length 2 cm. The device can be called a "projection chamber".

In the sketch of Figure 3, an high energy photon enters the bulk detector, a photoproduction

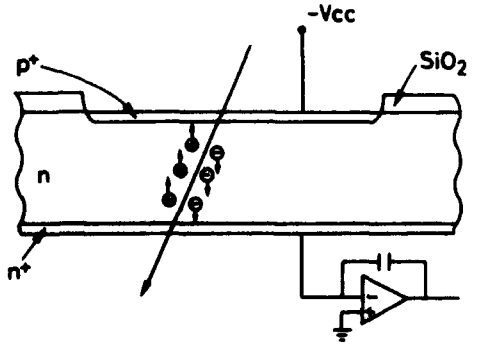


FIGURE 1
Schematic representation of a typical p-n junction silicon detector.

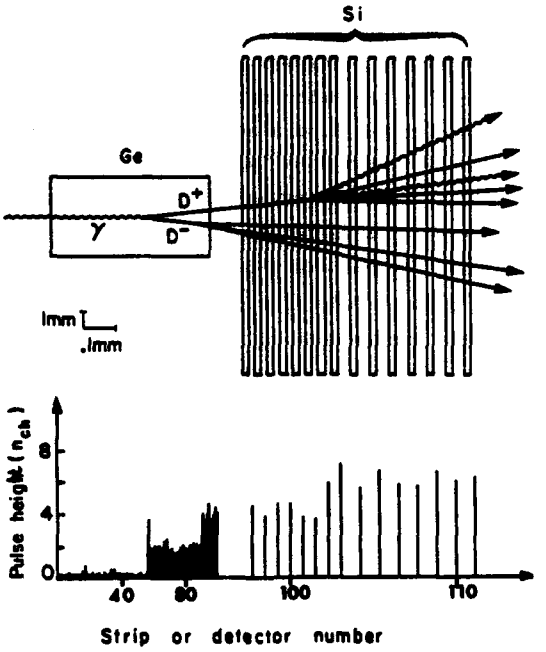


FIGURE 2
a) A set of p-n junction silicon detectors preceded by a Ge bulk detector where a D^+-D^- event occurs, and b) corresponding pulse height pattern at the detector's outputs.

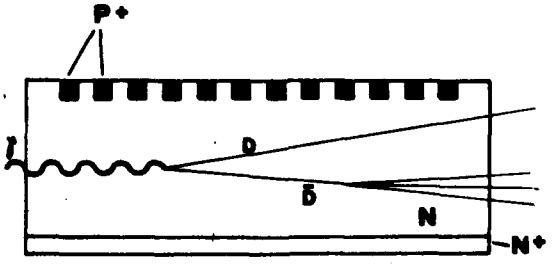


FIGURE 3
Ge bulk detector.

occurs with nucleous recoil and creation of minimum ionising particles.The ionisation density produced in the detector is sensed, as a function of the position along the axis, by the p^+ strips. The germanium bulk, detector , working at 90 K, has the advantage over the silicon telescope of a higher spatial resolution along the beam, 50-100μm vs 200μm.

Such a gain results from (a)a factor-two-higher specific energy loss in Ge than in Si, (b) the absence of air gaps between detection cells, (c) a better ENC resulting from the lower capacitance of the pick-up electrodes as strips instead of planes as in conventional detectors.

The device of Figure 3 is obtained from high purity n-type germanium. As the classical silicon planar technology can not be used with germanium, the definition of the pattern of the p^+ strips is obtained by selectively etching a continous boron implant using a Ti-Au film as masking material. The n^+ back contact is formed by lithium evaporation and diffusion[6,7].

3. SILICON MICROSTRIP DETECTOR

The first semiconductor device developed in order to get high spatial resolution is the silicon microstrip detector. Figure 4 shows a tipycal layout of a detector built by Kemmer, who pioneered the planar technology in high resistivity silicon[8].

The microstrip detector is essentially an array of a large number of reverse biased p-n diodes integrated on the same high resistivity wafer.

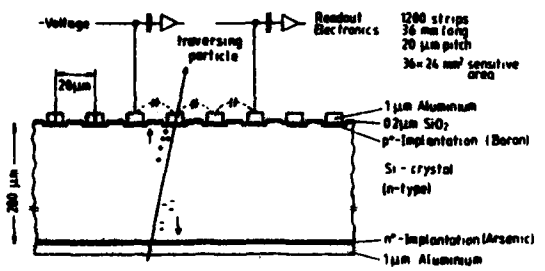


FIGURE 4
Cross-section of a microstrip detector with capacitative charge division.(From Nucl.Instr. and Meth. A235(1985)p.210)

The thickness of the wafer ranges usually from 200 to 300μm (the detector is intended for a particle incidence which is almost orthogonal to the surface of the wafer), the active area is of the order of 10-20 cm^2 and the pitch of the strips can be as low as 10-20μm.[9,10].

It is worthwhile to look in some details to the induced current pulses at the strips.

Let us consider a particular strip and a single electron-hole pair generated at point A near to the anode in front of the strip itself. In this particular case,the electron is immediately removed by the anode and doesn't induce any current on the considered strip. The hole travels in the depleted bulk until falls on the strip. The current induced at a given instant during its travel can be obtained knowing the velocity $u=\mu E(t)$ of the hole and a so called weighting field E_W [11,12]. E_W is the electrostatic field calculated giving potential 1 at the considered strip, 0 at all the remaining electrodes and considering the Si undoped.

Field lines of E_W are shown in Figure 5 together with the trajectory of the hole in the real field[13].

The current is given by

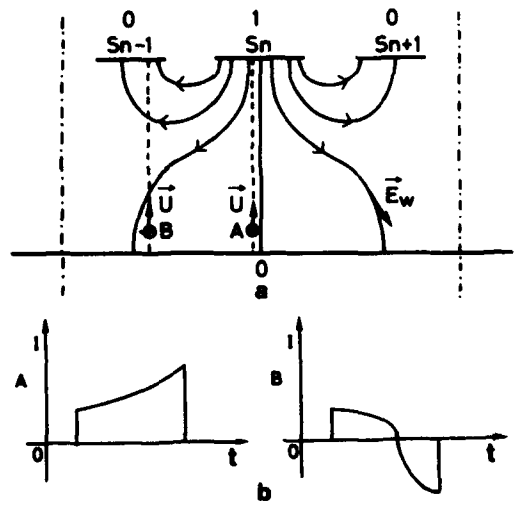


FIGURE 5
a) Schematic cross-section of a microstrip with the weighting field E_W highlightened. b) Induced current pulses at the S_n th strip for an electron-hole pair generated at point A or B.

(3) $I = -q \cdot u \cdot E_W$

It is apparent that u is always opposite to E_W for all the trajectory.
The current shape is shown in the figure and the total charge induced is

(4) $$\int I dt = -q \int u \cdot E_W \, dt = -q \int E_W \cdot dl$$

$$= -q(V_W(P_{in}) - V_W(S_n))$$

i.e.

(5) $$\int I dt = -q(0-1) = q$$

Let's now consider an hole starting at point B. During its travel along the drawn trajectory, it "sees" first a weighting field with component along the trajectory opposite to its velocity and then of the same sign as the velocity until it is collected by a side strip. The current shape given by (3) is also sketched in the figure and equation (4) gives

(6) $$\int I dt = -q(0-0) = 0$$

What has been shown for a hole starting at A or B can be easily generalized to an hole-electron pair generated at any point.
Generally speaking, charges moving in the depleted bulk induce current pulses in several strips, but only the strips collecting the charges give a signal with area different from zero, equal to the collected charge. Crosstalk between strips can therefore be avoided by suitable treatement of the signal.

The large number of strips and their small pitch would lead, in large area detectors, to formidable problems of fan-out, cost and power dissipation if a single amplifier should be connected to every strip.
Clever methods of readout have been developed to reduce connections. One is shown as an example in Figure 4. A charge amplifier is connected every three strips and, exploiting the interstrip capacitances, a weighted charge division of the charge collected by the strips occurs between the interested charge preamplifiers. Figure 6 shows a recent microstrip system with its fan-out and hybrid preamplifiers [14].



FIGURE 6
Picture of a silicon microstrip detector, glued on a ceramic motherboard, with kapton foils connecting the signal strips to the hybrid preamplifiers. (From Nucl.Instr. and Meth. 226 (1984) p.64)

For minimum ionizing particles generating 25000 electron-hole pairs within the detector, a resolution of 3μm RMS has been achieved, with a detector having 20μm pitch and one amplifier per strip. With the method of charge division the spatial resolution has worsened from 4.5μm, connecting one preamplifier every 3 strips, to 20μm for a preamplifier every 12 strips.
Integrated electronics (not on the same wafer of the microstrip detector) has already been designed and produced in NMOS tecnology. Single amplifiers are provided per each strip having pitch of 25μm. A spatial resolution better than 5μm RMS have been demonstrated [15,16].

4. CHARGE COUPLED DEVICES

Charge coupled devices (CCD) have been invented in 1970 at Bell Labs [17]. Commercially produced two dimensional CCDs, used for TV cameras, night vision systems, visible and Xray astronomy, have been considered as high precision detectors for ionising particles since 1981 by Damerell [18].

The CCD can be considered a fine matrix of potential wells just below the surface of the silicon, typically 20μm x 20μm x 10 μm depth.

Each one can store the charge delivered by an impinging particle and the image of hitted points is stored as charge in these pixels. The CCD is provided by a surface structure of channel stops and field electrode arrays which allows electrons from each pixel to be transferred from pixel to pixel and deposited in turn onto a single output collecting node to which a single on-the-chip preamplifier is connected.
The general layout of the device is shown in Figure 7[19]. Buses I1,I2 and I3 drive the horizontal field electrodes which transport the electrons stored in the pixel lines to a bottom linear CCD which receives in turn the electrons stored in each line. This latter CCD, driven by buses R1,R2 and R3, leads the electrons of each line sequentially to the output node.

As far as the interaction of the impinging particle with the detector is concerned, Figure 7 shows 3 regions of interaction. One, just underneath the surface, in which the generated electrons are completly trapped into the pixel potential well. A second, drawn as a dotted segment, in which electrons, generated in a field free region of high lifetime for minority carriers, diffuse freely and part of them eventualy fall in a potential well (not necessarily the nearest one). A third, corresponding to the p^+ substrate, where all electrons recombine. The effective thickness of the CCD for particle detection turns out to be approximately 15 μm.

This device has the obvious advantage of a single output channel, payed by a relatively long reading time of about 12ms.
Another advantage of the device is that, once a charge is stored in a pixel, it can no more spread out by diffusion so that spatial resolution it is not impaired by delayed readout.

Due to the small depth of a pixel, a small charge is deposited by a minimum ionizing particle (≈1250 electrons). However, working at 200 K and exploiting the advantages of low output capacitance and capacitance matching with the on-chip preamplifier, the very low ENC of 50 electrons has been reached, leading to a sufficient resulting signal-to-noise ratio.

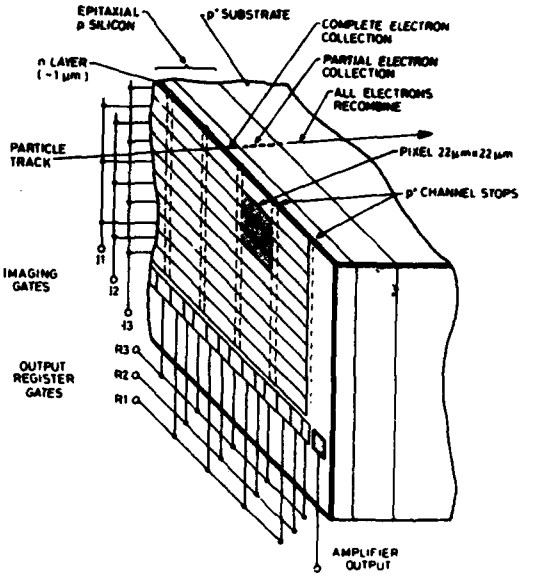


FIGURE 7
General view of a CCD detector showing the orthogonal arrangement of gates and channel stops defining the pixels. (From Nucl.Instr. and Meth. 213(1983) p.203)

In a recent experiment a clever way of partly overcome the long readout time has been devised [20,21]. The CCD has been kept continously running, with only a part of it exposed to the beam, the rest (the so called parking area) being shielded from the main radiation. Therefore the particle pattern registered at one end of the CCD move continously towards the opposite end. Once a trigger signal indicate the presence of an interesting event, the clock which provide the continous shifting, is stopped after a time which allows the recorded image to exit the region exposed to the beam and be freezed in the parking area, while the beam is turned off. The interesting image can be read at the most convenient rate at the output node not exposed to radiation of any kind.

These CCDs are powerful devices for recostructing vertices in elementary particles experiments. One example of reconstraction of charm dacays in the NA32 experiment running on the CERN SPS is given in Figure 8[20].
Two CCDs, placed at a distance of 10mm each other and 10mm from a copper target, enable the

accurate recostruction of a charm event by the identification of two vertices whith an high precision of the impact parameter (5μm). The two vertices are formed by tracks 3,4,6,7,8 and tracks 1,2,5 respectively. From the extimated distance of the two vertices the lifetime of the short living particle can be inferred.

It is interesting to note that CCD is a pixel device which attributes unambigously X,Y coordinates to each hit. On the contrary, a set of microstrip detectors with strips of different orientation define X,Y coordinates of the hits with the coordinates of hitted strips.When detectors are hit by a single event, coordinate attribution is unambigous also in this case,but

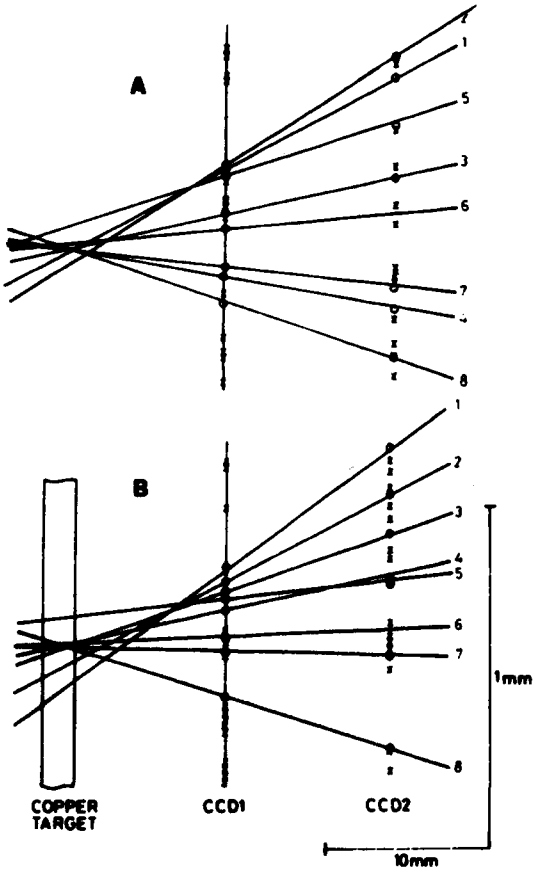
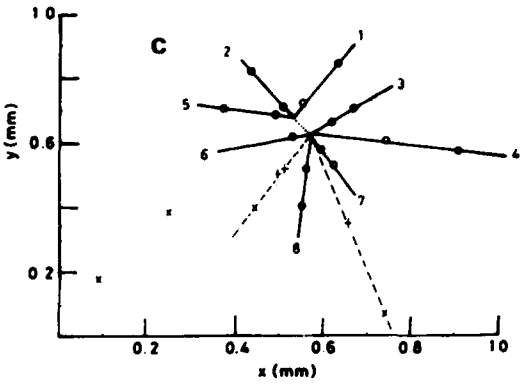


FIGURE 8
Reconstruction of a charm decay in the NA32 experiment at CERN SPS. a) and b) correspond to two different side views of the same event. c) is the projection along the beam direction. Primary and decay vertices show up unambigously.(From Nucl.Instr. and Meth. A253(1987) p.479)



for multiple events strips coordinates are not unambigously attributed to hits(ghost events).

5. SILICON DRIFT CHAMBER

In 1983 E.Gatti and P.Rehak introduced the "Solid State Drift Chamber",a completely depleted device with a low capacitance collecting anode[22]. The device has been then developed within an international collaboration involving Politecnico di Milano,Brookhaven National Lab., Technischen Univarsitat - Munchen, Max Planck Institut fur Physik und Astrophysik - Munchen and Lawrence Berkeley Lab[22-27].

The working principle of the device is best understood looking at the test structure of Figure 9. The figure shows how a thin n^-silicon wafer can be fully depleted by a peripheral n^+ anode and two p^+ electrodes placed on each surface of the wafer and held at the same potential by an external connection. As the reverse bias increases, the undepleted conductive layer in the middle of the wafer first narrows, then disappears and disconnects(from the ohmic point of view) the n^+ electrode with a dramatic drop of its capacitance with respect to the p^+ contacts, as shown in Figure 10[26].

A schematic view of the drift chamber is shown in Figure 11. Two arrays of p^+ electrodes on both surfaces of the wafer, biased at increasing potential energy, provide the drift field. A drawing of the potential energy in the drifting region is given in Figure 12 (for a

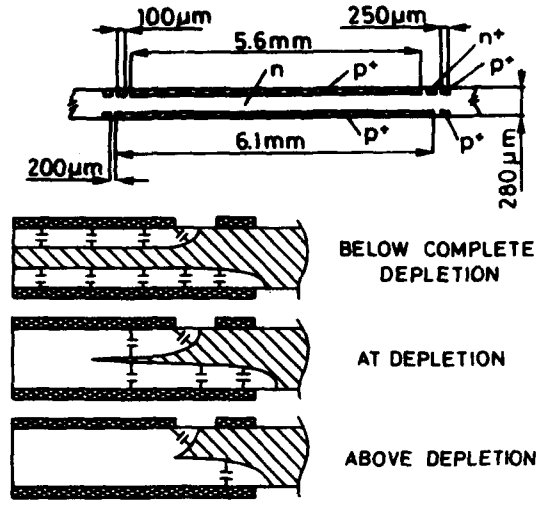


FIGURE 9
Silicon device to test the mechanism of complete depletion.

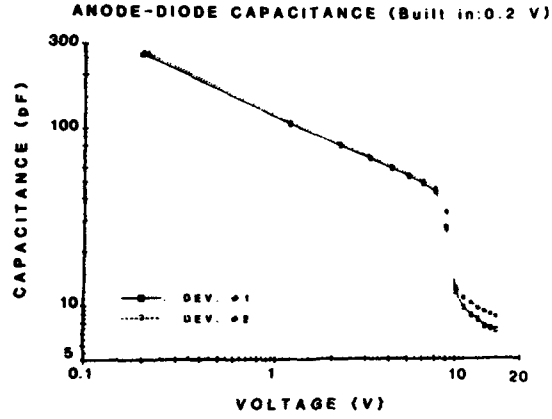


FIGURE 10
Capacitance between the two p^+ electrodes connected together and the small n^+ contact of Figure 9 as a function of the bias voltage.

voltage difference between strips of 15V and for a 10 KΩcm n-type completely depleted silicon). The electrons generated by the ionizing particle are focused at the bottom of the buried potential channel and drift towards the collecting anode, while holes, driven by the depletion field, are quickly collected by the nearest p^+ electrodes. Figure 13 shows how, by suitably biasing the electrodes at either side, it is possible to shift the bottom of the potential channel toward the surface where the collecting anode AR is placed.

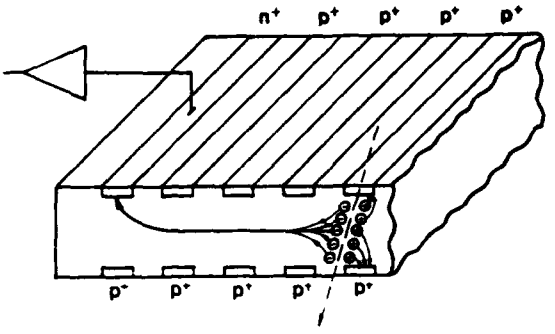


FIGURE 11
Schematic view of the drift chamber, showing the working principle.

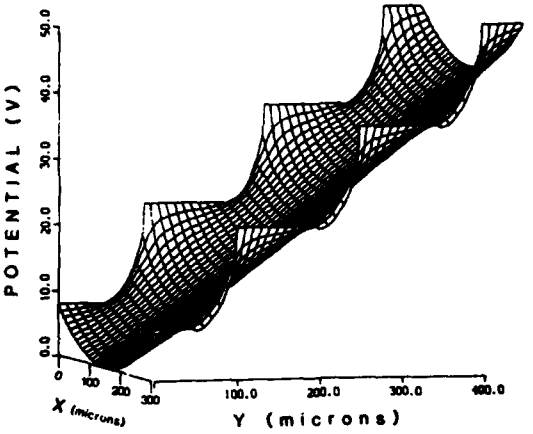


FIGURE 12
Potential energy for the electrons (negative electric potential) in the drifting region of the SDC.

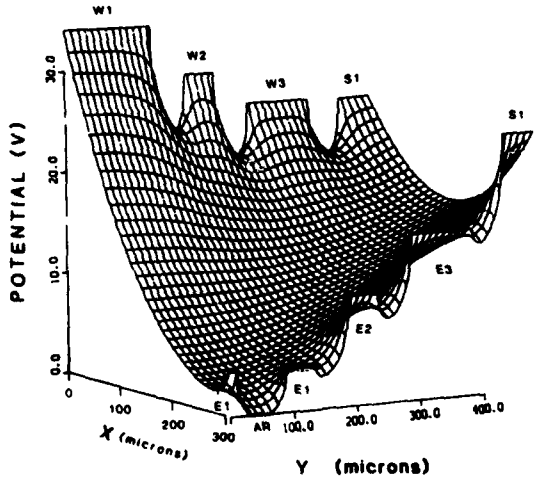


FIGURE 13
Potential energy for the electrons within the SDC close to the readout anode AR.

The cloud of electrons induces at the anode an output pulse only when the electrons arrive close to it because of the electrostatic shield of the p^+ electrodes, which behave as a Faraday cage. The drift time of electrons may be used to measure one of the interaction coordinates while the collected charge allows the measurement of the energy released by the incident ionizing particle.

Figure 14 shows how the pulses appear at the anode for different drift distances of the electron cloud in a structure like the one of Figure 11 [25].

Resolution is limited by the leakage current of the detector, series noise of the amplifier, Poisson noise due to the spread of the electron pulse because of diffusion [24]. The obtained resolution at room temperature for a wide range of drift fields is of the order of 4 μm RMS.

The drift chamber can be used, as the CCD, as a two-dimensional imaging device simply dividing the strip anode in pads. A 1x1.5cm^2 multianode SDC has been experimented [27] and a new large area (4x4 cm^2) is under developement.

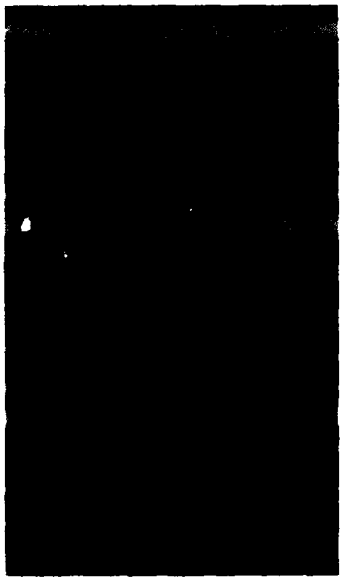


FIGURE 14
Signal pulses at the anode of a linear SDC for different drift distances of 1,2.5,3.85 mm respectively. Time scale is 200 ns/div.

The dinamics of electrons in drift chambers has been studied carefully with a two-dimensional chamber [28]. Figure 15 shows the expected electron distribution, for different time of flights, due to diffusion, Coulomb repulsion and the combined effects, for two different number of electrons in the cloud. Expected RMS width of the cloud has been experimentally checked by measuring how the charge is shared among the anodes. The results are reported in Figure 16. From the theory and the measurements it is apparent the importance of the Coulomb repulsion when the electron cloud is larger than 25000 electrons, which is just the charge released by a minimum ionizing particle traversing a depleted region of 300 μm.

Diffusion does not pose strong limitations to two-dimensional SDCs as far as resolution is concerned and may even be beneficial along the anode's coordinate because allows centroid evaluation based on charge shared among contigous anodes. On the other hand, it is detrimental as far as recognition of close double-hits is concerned. Typical resolution for double hit is estimated in 300 μm. However, estimation of the presence of a double hit can be assessed with a great degree of confidence from the measurement of the released charge.

Similarly to CCD, the coordinates of the hits are determined without ambiguity for a two-dimensional SDC because each electron cloud has one coordinate attributed according to the interested anode and the other according to the timing of its pulse.

As far as speed of readout is concerned, SDC is intermediate between completely parallel readout of microstrip detectors and completely serial readout of CCDs. Typical readout time for SDC is 1-2 μs per centimeter of drift distance.

Generally, in high energy physics experiments, start time of the interesting event is available from an indipendent trigger pulse and so timing of pulses at the anode is sufficient for drift time measurement. However in Xray or

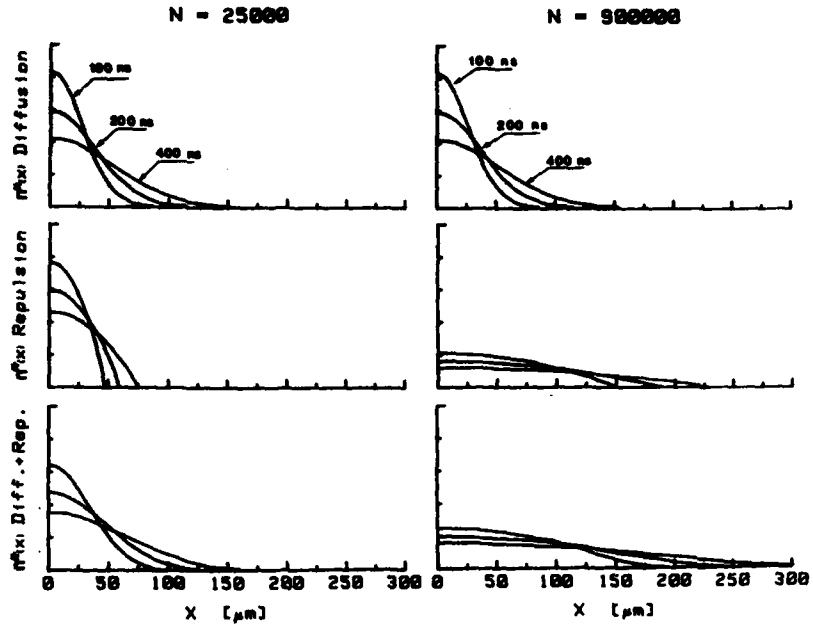


FIGURE 15
Density distribution of electrons along the drift direction at different times. Distributions as due to diffusion only, mutual repulsion only and the combined effects are plotted for two different numbers of electrons in the cloud.

γ spectroscopy, an independent time information is not available from other detectors. A prompt signal from the interaction can be obtained in the SDC from all the field electrodes of one side of the wafer, paralleled by capacitors and connected to a charge amplifier[27]. As the interaction is localized, a prompt pulse is induced both by the electrons falling into the bottom of the valley and by the holes traveling toward the field electrodes. Figure 17 shows such pulse which has opposite sign depending in which half of the detector the Xray has been absorbed. The charge amplifier connected to the anode is insensitive to the interaction because the movement of charges is shielded by the field electrodes until electrons leave the drift region and approach the anode. At this moment a delayed pulse is detected at the anode and one of opposite sign at the paralleled field electrodes.

Surface effects must be carefully considered in designing SDC. In particular it has been observed that a surface hole current is exchanged between two neighbour p^+ strips of the depleted detector when a potential difference is applied between them. It is negligible until a given threshold voltage is reached, then it rises very sharply. This effect can be explained by considering the potential energy close to the surface of the detector under the oxide between contiguous electrodes, as shown in Figure 12 and

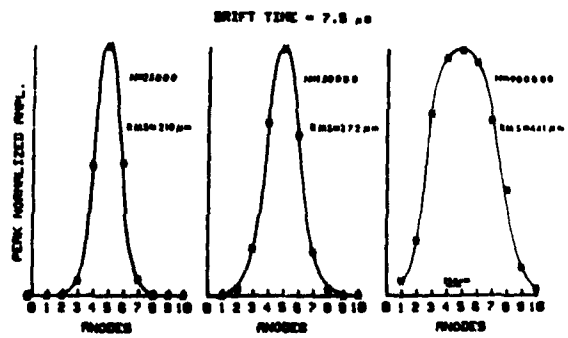


FIGURE 16
Measured normalized charge at the anodes for different N and a drift time of 7.5 s.



FIGURE 17
Output waveform of a self-timing circuit. Central trace: anode output. Upper (lower) trace: photon incident on the top (bottom) side of the detector (time scale 200 ns/div).

13. A potential barrier close to the surface limits the flow of holes from the electrode at lower potential energy toward the contiguous one at higher energy. As the voltage difference between the two strips grows, this potential barrier decreases and consequently the flow of holes increases rapidly.The hole barrier almost disappears at the mentioned thereshold voltage between strips and, in this condition, the hole current would diverge if not limited by the external circuit.

A maximum potential difference can be therefore applied between neighbour strips. This maximum potential difference is a function of the distance between the two strips and depends on the properties of the oxide between them. In fact, positive oxide charges induce, as it is well known, a sheet of mobile electrons at the semiconductor-oxide interface , which in turn influence the heigth of the barrier for the holes flow. Environment and biasing conditions affect this electron sheet, and in turn the stability of the threshold voltage for the hole flow (in Figures 12 and 13 this effect is not taken into account).

As a consequence of these phenomena, there is a practical limit to the drift field in the SDC. In fact, while the increase of the gap between strips increases the potential barrier to the flow of holes, on the other hand it also increases the electric field on the edge of the p^+ strip at higher potential energy (breakdown problems) and increases the slope perturbation at the bottom of the drift channel. Practical values so far reached are of the order of 1000V/cm.

It is important to note that this flow of holes does not affect the anode leakage current, as shown clearly in Figure 18.

The exchange of holes between p^+ strips can be useful to self bias arrays of field electrodes in SDC. In fact, given an array of p^+ strips over a depleted wafer, if a current is applied between the first and the last strip, a hole current flows from the first contact to the last one passing through all the intermediate p^+ floating electrodes and the resulting voltage difference is nearly equally devided between all contiguous strips. This effect has been called "Chain effect" and the exchanged hole current "Chain current".

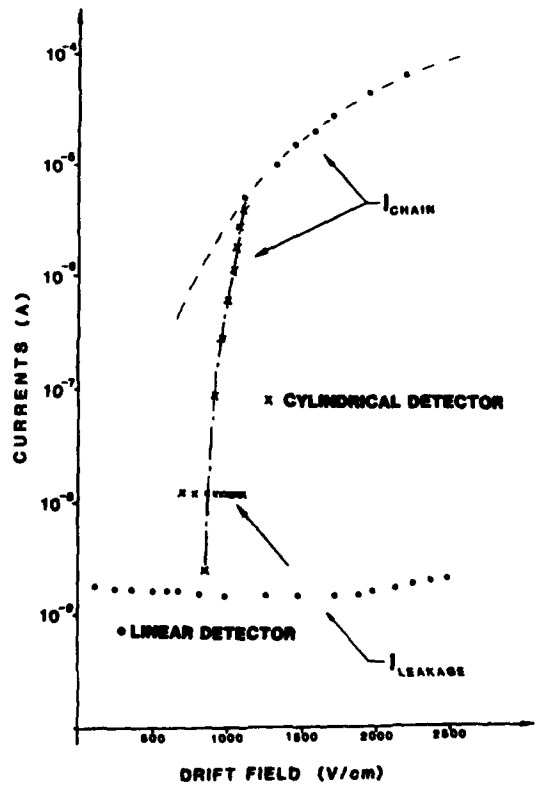


FIGURE 18
Anode and chain current for a linear SDC and a cylindrical SDC vs drift field.

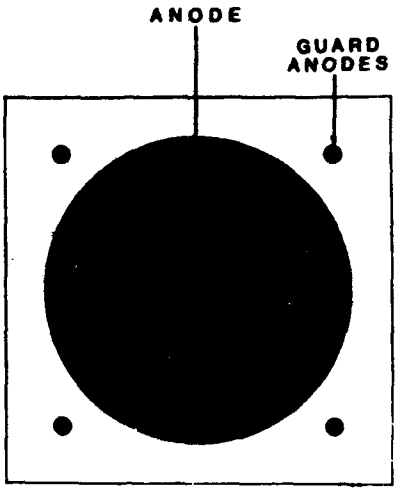


FIGURE 19
Cylindrical drift chamber.

A drift detector intended mainly for energy measurements is shown in Figure 19[27]. The p^+ field electrodes are concentric rings on both faces of the wafer whith a small n^+ collecting electrode at the center of the detector on the upper face only.The drift field has a cylindrical symmetry and the electrons drift radially toward the collecting electrode. The active area of the detector has a diameter of 1cm and the output capacitance is only 0.06pF due to the very small dimensions of the collecting electrode (200μm of diameter) surrounded by a completely depleted silicon.The leakage current at room temperature was about 3nA.

The p+ circular electrodes are "self biased" by the chain effect. A voltage difference is applied between the first and the more external ring.The injection of electrons from the external undepleted n region toward the central anode is avoided by biasing the undepleted region at a potential energy lower than that of the last external p^+ ring which defines the active volume of the detector.

Figure 20 shows the spectrum of the ^{241}Am measured at room temperature with a shaping pseudo-Gaussian and a peaking time of 250ns[27]. Equivalent noise charge was only 110 electrons. Note that the detector capacitance was only 1% of the total input capacitance. Further improvement would therefore be possible by integrating the input transistor on the detector. Measurement performed at liquid nitrogen temperature showed 36 electrons of equivalent noise charge.



FIGURE 20
^{241}Am spectrum obtained with a cylindrical drift detector at room temperature (linear scale). The highest peak is the L at 17.8 KeV, while the right one is the line at 59.54 KeV.

Low output capacitance drift detectors with field electrodes implanted only on one side of the wafer have been recently produced by Kemmer [29].

REFERENCES

1 K.McKay, Phys.Rev. 84 (1951) 829.

2 E.Gatti e P.F.Manfredi, Riv. Nuovo Cimento 9, N.1 (1986)

3 V.Radeka, Nucl. Instr. and Meth. 226 (1984) 209

4 V.Radeka, Nucl. Instr. and Meth. A253 (1987) 309

5 S.Amendolia et al., Nucl. Instr. and Meth. 226 (1984) 78

6 S.Amendolia et al., IEEE Trans. on Nucl. Sci. 31 (1984) 945

7 S.Amendolia et al., IEEE Trans. on Nucl. Sci. 30 (1983) 98

8 J.Kemmer, Nucl. Instr. and Meth. 169 (1980) 449

9 R.Klanner, Nucl. Instr. and Meth. A235 (1985) 209

10 B.Hyams et al., Nucl. Instr. and Meth. 205 (1983) 99

11 S.Ramo, Proc. IRE 27 (1939) 584

12 G.Cavalleri et al., Nucl. Instr. and Meth. 21 (1963) 177

13 U.Kotz et al., Nucl. Instr. and Meth. 235 (1985) 481

14 H.M.Heijne et al., Nucl. Instr. and Meth. 226 (1984) 63

15 J.Walker et al., Nucl. Instr. and Meth. 226 (1984) 200

16 G.Anzivino et al., Nucl. Instr. and Meth. 243 (1986) 153

17 W.S.Boyle and G.E.Smith, Bell System Technical Journal 49 (1970) 587

18 C.J.S.Damerell at al., Nucl. Instr. and Meth. 185 (1981) 33

19 R.Bailey et al., Nucl. Instr. and Meth. 213 (1983) 201

20 C.J.S.Damerell et al., Nucl. Instr. and Meth. 253 (1987) 478

21 C.J.S.Damerell, Lectures presented at the Advanced Study Institute on Techniques and Concepts of High Energy Physics,St.Croix, june 1986.

22 E.Gatti and P.Rehak, Proc. 1983 DPF Workshop on Collider Detectors, LBL-15973 UC-37, Conf. 830224 (1983) 97

23 E.Gatti and P.Rehak, Nucl. Instr. and Meth. 225 (1984) 608

24 E.Gatti et al., Nucl. Instr. and Meth. 226 (1984) 129

25 E.Gatti et al., Nucl. Instr. and Meth. 235 (1985) 224

26 E.Gatti et al., IEEE Trans. on Nucl. Sci. NS-32 (1985) 1204

27 P.Rehak et al., Nucl. Instr. and Meth. A248 (1986) 367

28 E.Gatti et al., Nucl. Instr. and Meth. A253 (1987) 393

29 J.Kemmer and G.Lutz, Nucl. Instr. and Meth. A253 (1987) 365

Session A1.1

Bipolar-CMOS

Chairman: G. Zocchi

Monday, September 14, 1987

THE USEFULNESS OF ADVANCED DRAIN STRUCTURES AS EMITTERS IN SCALED BICMOS

J. Winnerl, F. Neppl, B. Vollmer, M. Stegherr, B. Pfäffel

Siemens AG, Central Research and Development,
Microelectronics, Otto-Hahn-Ring 6,
D-8000 München, FRG

Bipolar transistors can be implemented in a CMOS technology without excessive expense using the n-channel source/drain implantation simultaneously for the formation of the bipolar emitter. During MOS scaling the change of the drain structures from phosphorus to DID and LDD influences the individual bipolar device parameters. However this BICMOS concept can be extended to sub-µm CMOS without loosing performance.

1. INTRODUCTION

CMOS performance can be improved considerably by utilizing the higher transconductance, the threshold stability and the higher speed and current drive capability of bipolar transistors in analog and digital circuits. These optional bipolar transistors should be implemented in a CMOS technology with little extra expense and without affecting the MOS device characteristics. The possibility to use the n-channel S/D implantation for the emitter formation is especially investigated. Scaling CMOS implies changes of the n-channel drains from phosphorus to DID (double implanted drain) and LDD (lightly doped drain) to alleviate hot carrier effects. To clearify the usefulness of this concept for scaled BICMOS the impact of different drain structures on the bipolar performance is studied.

2. PROCESS CONCEPT

The BICMOS technology presented here is based on an n-well CMOS technology (Fig.1). This allows to realize isolated npn bipolar transistors with the n-wells used as collectors. The emitters are simultaneously implanted with the n-channel source/drains so that no additional photolithography or other process steps are necessary for emitter formation. As n-channel source/drain configurations phosphorus drains, DID, LDD were used which are typical for 2µm, 1.5µm and 1µm CMOS, respectively.

The base was selectively implanted in the bipolar transistor. This step requires the only absolutely necessary additional photomask to realize isolated bipolar transistors in CMOS. The base implantation was adjusted for the different emitter complexes.

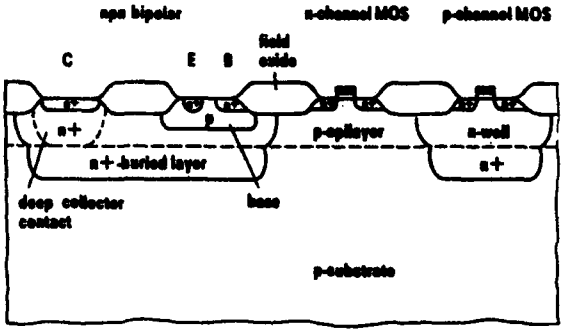


FIGURE 1

Cross section of an n-well CMOS structure with an isolated npn bipolar transistor.

For applications that need a high current drive capability like buffer circuits a buried layer and a deep collector contact were added requiring two extra masks.

3. BIPOLAR DEVICE STRUCTURE

During CMOS scaling the n-channel drain structure is improved with respect to hot carier and short channel effects. While phosphorus drains were typical for 2µm CMOS DID (double implanted drain) and LDD (lightly doped drain) are used for 1.5µm and 1µm CMOS and beyond. This implied changes in bipolar emitter structure, too. For phosphorus drains phosphorus emitters with junction depths of about 0.5µm are obtained (Fig.2a). For the DID and LDD arsenic emitters can be obtained without additional photolithography (Fig.2b and c). For LDD as an option the low dose of phosphorus can be implanted unmasked. In this case the optimizations of the n- and p-channel MOSFETs are coupled and the bipolar transistor has an arsenic emitter with phosphorus tail (Fig.2c).

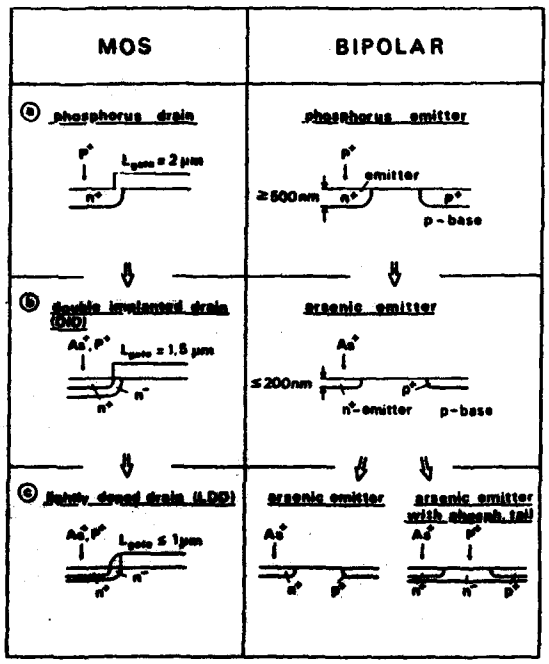


FIGURE 2

Emitter structures obtained simultaneously with n-channel source/drain implantations in scaled BICMOS.

LDD structures arsenic and phosphorus are implanted in the n-channel regions while the p-channel regions are usually masked. With this mask the phosphorus implantation can also be avoided in the emitter regions. Thus for DID

4. BIPOLAR DEVICE CHARACTERISTICS

For phosphorus a relatively deep emitter with about 0.5µm emitter junction depth is obtained. For DID and LDD n-channel drains with masked low dose phosphorus implantation the arsenic emitters are as shallow as 200 nm. The base implantation dose and energy had to be readjusted for comparable base pinch resistivity. It was necessary to account for the reduced emitter junction depth and the lacking emitter push effect inherent in phosphorus emitters. The base dose had to be reduced by a factor of 4 for the arsenic emitter. As seen

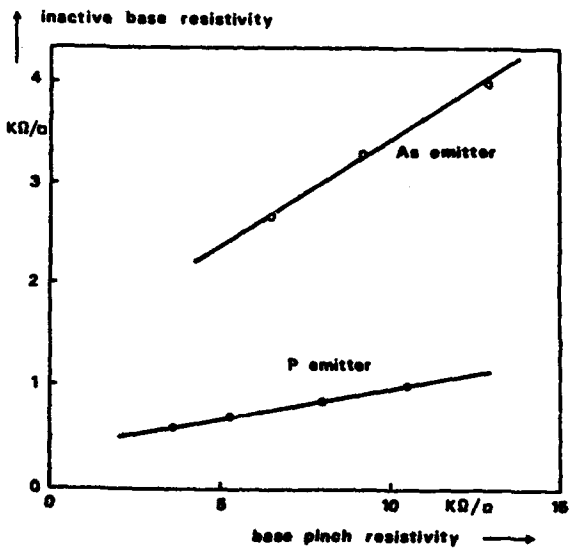


FIGURE 3

Resistivity of the inactive base depending on the base pinch resistivity for P and As emitters.

from Fig. 3 this results in an increased resistivity of the inactive base in a similar amount. This inactive base resistivity strongly determines the base series resistance which in turn affects the dynamic behaviour of the transistor with As emitter. Since for As emitters there is less compensation of base charge, however, there is less sensitivity to process fluctuations.

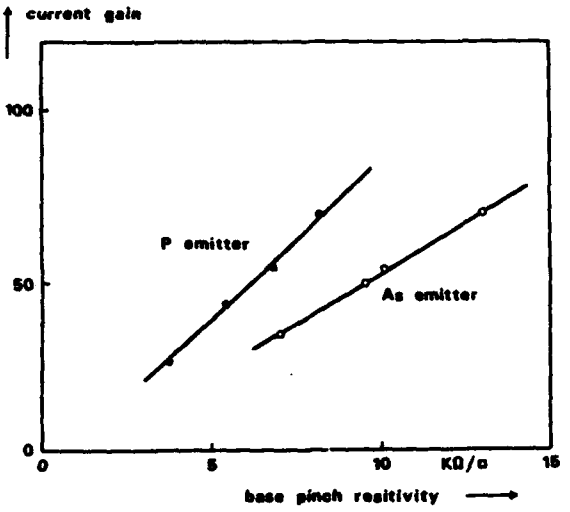


FIGURE 4

Bipolar current gain as a function of the base pinch resistivity for P and As emitters.

Fig. 4 shows that for a given base pinch resistivity transistors with As emitters have lower current gains, which is due to the worse emitter efficiency of the shallower As emitter. The reason for this is the vicinity of the emitter contact to the emitter base junction, which incrases the effective hole recombination in the emitter and thus the hole diffusion current into the emitter.

The cutoff frequency is typically 2.5 GHz for As emitters compared with 1.8 GHz for P emitters.

5. CIRCUIT PERFORMANCE

To study the influence of the changed device structures on the BICMOS performance, typical analog and digital circuits are investigated. In Fig. 5 a differential amplifier with 50Ω

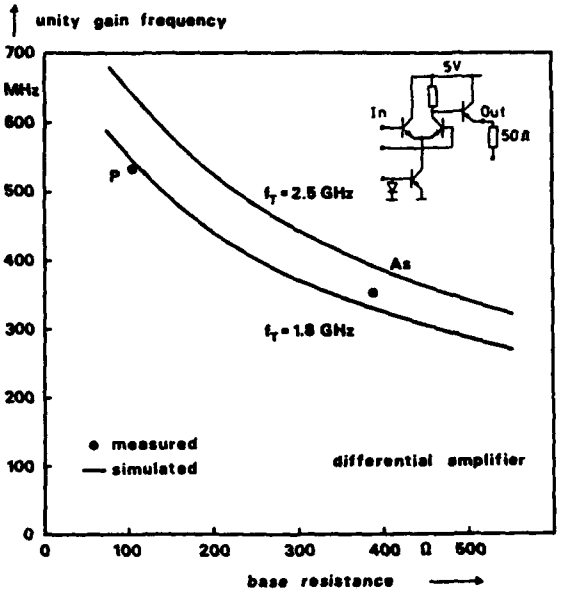


FIGURE 5

Measured and simulated unity gain frequency of a differential amplifier depending on cutoff frequency and base resistance.

load is characterized by its unity gain frequenzy. A comparison of the measured data shows a lower unity gain for As emitters than for P ones in spite of their higher cutoff frequency. Using circuit simulations the influence of the base resistance and the cutoff frequency were separated. The speed improvement by the higher cutoff frequency for As is shown to be even overcompensated by the increased base resistance. In Fig. 6 the influence of the current gain is shown. When a threshold current gain of about 50 is exceeded a further increase does no longer improve the speed performance.

In Fig. 7 measured an simulated gate delays are compared for an ECL inverter ring oscillator as a typical digital application. Similar as for the differential amplifier the improvement by cutoff frequency is compensated by the increased base resistance. The As sample exibits a 10% higher gate delay compared with P. Looking at a more complex circuit the maximum operating frequency of a static ECL frequency divider shows only a marginal difference between 770 MHz for P and 780 MHz for As. In this case the speed limiting influence of the base resistance seems to be not so pronounced than for the ECL inverter.

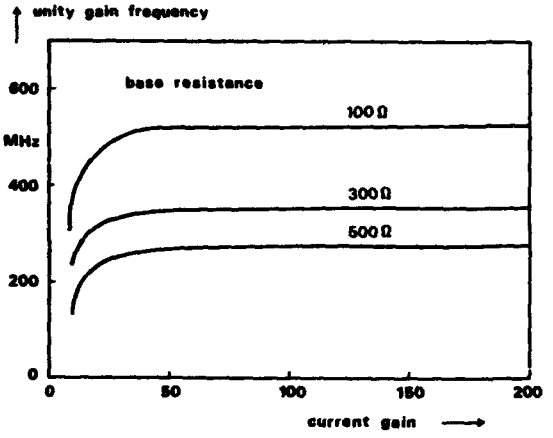


FIGURE 6 .

Dependence of the unity gain frequency of a differential amplifier on the bipolar current gain.

For the As sample base dose and energy are not yet optimized. A fine tuning of the base implantation is expected to further increase the cutoff frequency and to reduce the base resistance so that the small loss in speed for As is eliminated.

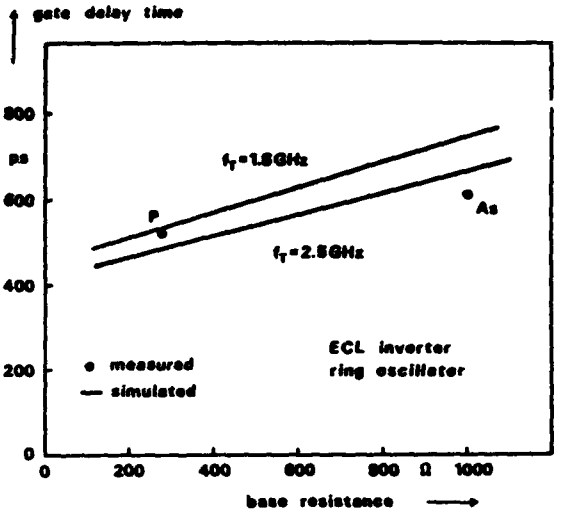


FIGURE 7

Measured and simulated gate delay time of an ECL inverter ring oscillator depending on cutoff frequency and base resistance.

6. CONCLUSIONS

It was shown that the emitter type influences the individual bipolar device parameters. The performance of bipolar transistors in BICMOS is not represented by a single parameter like the cutoff frequency alone. The combination of all essential parameters has to be considered.

For the applications considered the consept of using the MOS source/drains simultaneously as bipolar emitters can be extended to sub-µm CMOS.

ACKNOWLEDGEMENTS

This work has been partly supported by the European Community. The authors alone are responsible for the contents.

1.2μm Bi-CMOS Technology with High Performance ECL

H. IWAI, Y. NIITSU, G. SASAKI, M. NORISHIMA, K. SHINO*, Y. UNNO**,
K. TSUGARU*, H. HARA*, Y. SUGIMOTO*, and K. KANZAKI

Semiconductor Device Eng. Lab., Toshiba Corp., 1, Komukai-Toshiba-cho, Saiwai-ku, Kawasaki, 210, Japan. * Semiconductor Group., Toshiba Corp, ** Toshiba Micro-computer Engineering Corp.

1.2μm Bi-CMOS technology with ECL gate for high speed device has been developed. A process is carefully optimized for obtaining the best performance of ECL gate without degrading 1.2μm Bi-CMOS performance and mass productivity.

1. INTRODUCTION

Today, Bi-CMOS technology is being introduced in various kinds of high speed VLSIs. In these device, as well as Bi-CMOS gates, usually TTL or ECL gates are used in the I/O circuits to ensure the compatibility and the drivability of I/O interface. But, until very recently[1], there have been very few reports that ECL gates are used for internal logic in the Bi-CMOS VLSIs.

Although, ECL is currently the fastest LSI commercially available, its huge power consumption limits the number of gates integrated in a chip. Therefore, for the highest speed VLSIs and ULSIs, Bi-CMOS with internal logic ECL is the most realistic candidate, though level conversion circuit delay between CMOS/Bi-CMOS and ECL gates should be treated carefully. In this type of devices, it is possible that the three types of gates can be used appropriately according to the purpose, i.e.; CMOS gates for small load, high integration and low power; Bi-CMOS gates for medium load and medium integration; ECL gates for high speed or/and large load and small integration.

In these devices, it is desirable that process is optimized for all of the three types of gate. Several process possibilities are considered, and in our 1.2μm ECL/Bi-CMOS, process is firstly optimized for CMOS and Bi-CMOS gates and then for ECL gates. The results show that relatively high performance ECL gate is obtained, even when ECL gates process is optimized within 1.2μm Bi-CMOS technology. In this paper, 1.2μm Bi-CMOS technology with relatively high performance ECL is explained.

2. PROCESS

Fig.1 shows the cross section of 1.2μm Bi-CMOS structure. Fig.2 shows the outline of the process sequence. Table 1 shows main process parameters.

Basic process is 1.2μm CMOS technology where twin well is used for high integration. LDD structure is used for NMOSFET to ensure high reliability. For PMOSFET, conventional structure is used, and minimum poly Si gate length is 1.4μm to keep the effective channel length to the same value with that of NMOSFET.

There are several ways to combine CMOS and bipolar process. In our case, 2nd ploy Si emitter structure is chosen , and a few bipolar process steps boxed in Fig.2 are added to the CMOS process, which results in 5 mask step increase. But in the case of static RAM which alredy have 2nd poly Si process, only 3 mask step increase is required. The remainder portion of bipolar process steps is performed in the course of the original CMOS process. For example, collector N^- region is formed by N well diffusion, and base contact P^+ region is formed by source and drain ion-implantation of PMOSFET.

Buried N^+ and deep N^+ regions are necessary to reduce collector resistance and thus to obtain high f_T of bipolar transistor. Buried P^+ region is optional dependent on isolation rule, and is not used in this case. P type epitaxial layer is grown after buried N^+ formation. Isolation of bipolar transistor is performed by the P well.

In this process, 2nd poly Si emitter structure is chosen to reduce the emitter size and junction depth, which results in the high performance of bipolar transistor. In order to reduce the junction depth, base and emitter should be formed completely after CMOSFET fabrication. 2nd poly Si emitter structure is suitable for this kind of process sequence as shown in Fig.2. 1st poly emitter structure is not applied here, in which case base is doped before poly Si gate deposition and base width becomes as large as 0.3μm.

Fig.3 shows a typical impurity profile of bipolar transistor. To optimize the base resistance and h_{fe} separately, base ion-implantation is performed twice. Shallow and higher concentration peak is made by BF_2 implantation to form relatively high concentration P type base layer and to reduce the base

resistance. Deep and lower peak is made by B implantation to form intrinsic base P^- region and to control h_{fe}. The shallow peak has to be within the emitter region with enough margin so that it does not affect h_{fe}. For this reason, emitter depth is made a little deeper (around 0.15μm) within the succeeding short 900°C heat process by doping poly Si with a small amount of phosphorus in addition to arsenic.

To achieve low emitter resistance, emitter poly Si is deposited without native oxide layer at the interface [2]. One of the merit of poly Si emitter is to gain higher h_{fe} due to the existence of native oxide at the poly and single Si interface. But as emitter size reduces, resistance due to native oxide becomes significant. It was said that, when native oxide thickness is less than 10 Å, emitter resistance is allowerable [3], but control of native oxide thickness is not easy and variation of native oxide thickness sometimes causes significant h_{fe} variation. Therefore, we avoided native oxide during poly Si deposition. For higher heat process such as 1000°C, native oxide layer breaks up, but in 900°C process it is difficult to break up once native oxide layer exists.

DEVICE CHARACTERISTICS

Fig.4 shows the I_d - V_d characteristics of bipolar, NMOS and PMOS transistor. Table 2 shows main device parameters. CMOS characteristics is the same as those fabricated by original 1.2μm pure CMOS process. For bipolar transistor, f_{Tmax} of 6 GHz is obtained. Variation of h_{fe} in the wafer is very small.

Fig.5 shows a typical output waveform of 31 stage ECL gate ring oscillator. Fig.6 shows ECL gate propagation delay time versus sheet resistance of high concentration P type base layer. Sheet resistance of 1kΩ corresponds to r_{bb}' of 200Ω. t_{pd} of 0.17 nsec is obtained when ρ_s is 1kΩ. r_{bb}' reduction is important for ECL gate performance. Fig.7 shows propagation delay time versus load capacitance of 2 input nand Bi-CMOS gate. t_{pd} of 0.5 nsec is obtained when C_L is 0.7pF.

Fig.8 shows f_T dependence of τ_{pd} of Bi-CMOS gate by SPICE simulation. Though higher f_T is desirable, τ_{pd} of Bi-CMOS gate tends to saturate above 5-7 GHz, because performance of CMOSFET limits that of Bi-CMOS gate. Similarly, further improvement of other device parameters such as resistance and capacitance do not contribute to the improvement of Bi-CMOS gate performance significantly. For example, even when r_{bb}' value becomes half or twice, t_{pd} of Bi-CMOS changes only a few percent at the load capacitance of 0.7 pF. In conclusion, although further improvement of device parameters are desirable, as far as 1.2μm Bi-CMOS gate is concerned, f_T of 5-7 GHz is enough and base resistance is not so critical.

Situation is different for ECL gate. Both highest f_T and lowest r_{bb}' is necessary to obtain the best performance [4]. The base resistance can be optimized for ECL, but further improvement in f_T is not easy by the conventional bipolar structure of the 1.2μm Bi-CMOS process. For further improvement, such as walled emitter or emitter-base self alignment structure is desirable, which is currently not cost-effective and not technologically easy in order to combine with matured and mass-productive 1.2μm CMOS process. But in near future these advanced bipolar structure would be necessary for high performance Bi-CMOS with ECL circuits.

SUMMARY

1.2μm Bi-CMOS technology with relatively high performance ECL circuits has been developed. 2nd poly Si emitter structure is used to reduce emitter size. Poly Si is deposited without interface native oxide. Junction depth of bipolar transistor is reduced by forming the base and emitter regions completely after CMOSFETs fabrication. Process is optimized for ECL within the 1.2μm Bi-CMOS process, specifically for the base resistance. t_{pd} of 0.17 nsec is obtained for ECL gates. For further improvement of ECL gate performance, some advanced bipolar structure would be required. Bi-CMOS technology with internal ECL gate will be the most realistic candidate for the highest speed VLSI and ULSIs.

ACKNOWLEDGEMENT

The authors are grateful to Y. Unno, K. Namimoto, and H. Yamada for encouragement and useful suggestions about the work. They also thank M. Taguchi, T. Koyanagi, Aoyama, M. Kimura, and Watanabe for helpful discussions and support for fabrication.

REFERENCES

[1] F. Omerod, D. W. Schucker, K. Deierling, and N. Salamina, "A mixed technology gate array with ECL and BiMOS logic on a single chip", Dig. Tech. Papers, VLSI Symp. Circuit (Karuizawa, Japan), pp.31-32, May,1987.

[2] To be published

[3] A. W. Wieder, "Submicron bipolar technology - new chances for high speed applications", IEDM Tech. Dig., pp.8-11, December 1986.

[4] T. Nakamura, K. Ikeda, K. Nakazato, K. Washio, M. Namba, and T. Hashida, "63 ps ECL circuits using advanced SICOS technology", IEDM Tech. Dig., pp.472-475, December 1986.

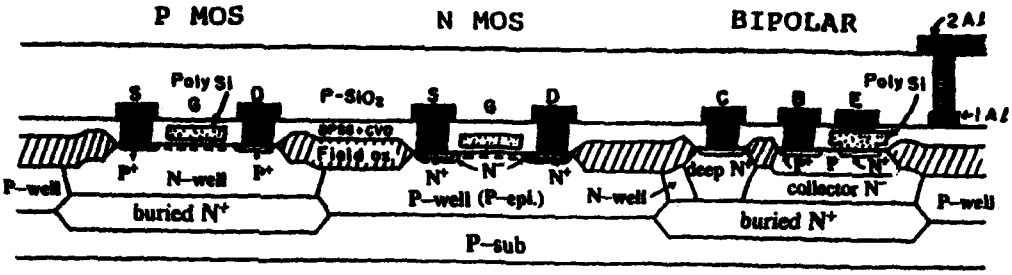


Fig.1 Cross-section of 1.2μm 2nd poly-Si emitter Bi-CMOS

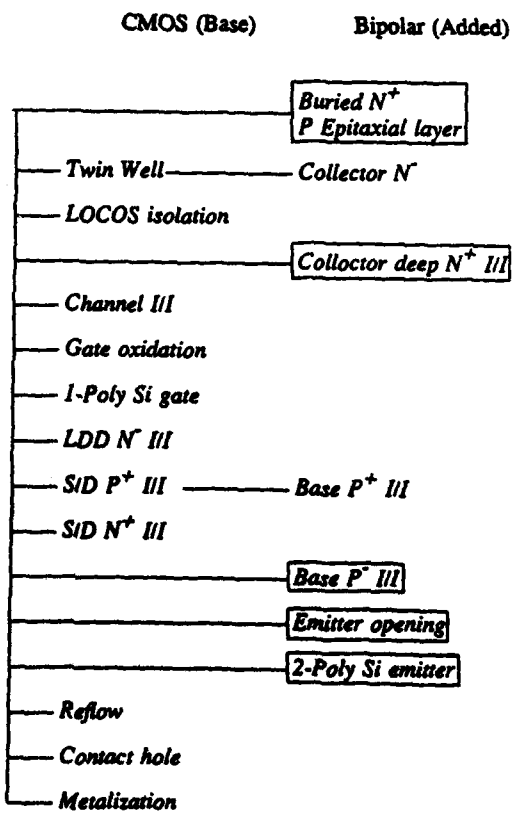


Fig.2 Outline of process sequence

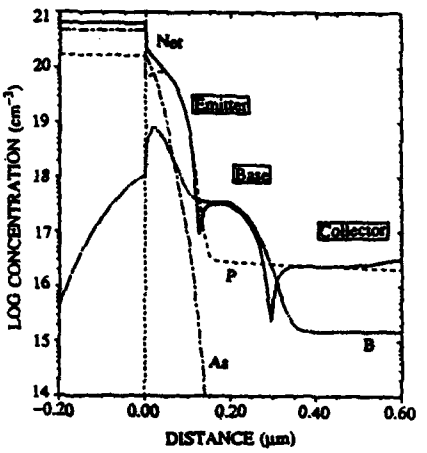


Fig.3 Impurity profile of bipolar Transistor

CMOS		
Twin well		
t_{ox}	250 Å	
N ch:	LDD	
	L_{poly}	1.2μm
	x_{jN}	0.18μm
P ch:	Conventional	
	L_{poly}	1.4μm
	x_{jP}	0.35μm
Bipolar		
Emitter Width		1.4μm
$x_{j\ EMITTER}$		0.15μm
$x_{j\ BASE}$		0.30μm
$t_{EPITAXIAL}$		2μm

Table.1 Main process parameter of Bi-CMOS

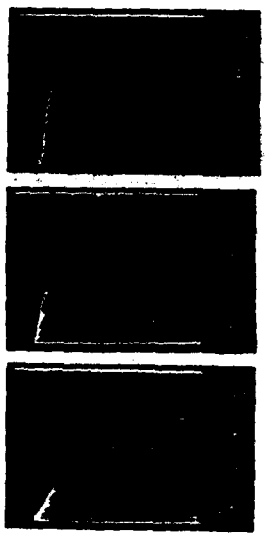


Fig.4 I_d - V_d characteristics of Bi-CMOS devices

CMOS

N ch:	V_{th}	0.8V
P ch:	V_{th}	−0.8V

Bipolar

h_{fe}	100
f_{Tmax}	6GHz
$r_{ee'}$	20Ω
$r_{bb'}$	200Ω
$r_{cc'}$	100Ω
C_{eb}	25fF
C_{bc}	30fF
C_{cs}	80fF
S_E	1.4μm × 5μm

Table.2 Main device parameter of Bi-CMOS

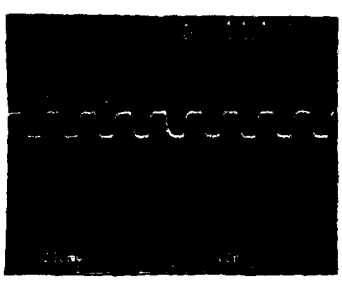


Fig.5 Typical output waveform of 31 stage ECL ring oscillator

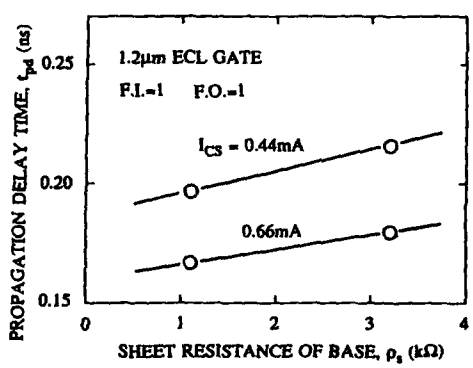


Fig.6 ECL gate performance
(t_{pd} versus ρ_s of high concentration P type base region)

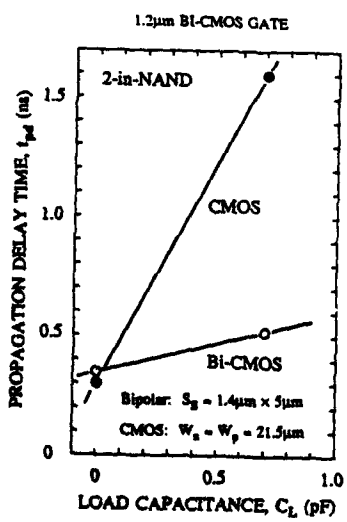


Fig.7 Bi-CMOS gate performance (t_{pd} versus C_L)

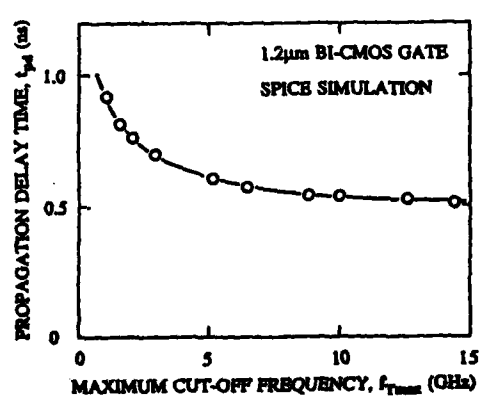


Fig.8 f_T versus t_{pd} of Bi-CMOS gate (C_L = 0.7 pF)

MULTIPOWER BCD 250V : A VERSATILE TECHNOLOGY TO REALIZE HIGH PERFORMANCE PICs

A.Andreini C.Contiero P.Galbiati

SGS Microelettronica SpA, Monolithic Microsystem Division
20019 Castelletto di Settimo Milanese, Milan, Italy

The feasibility of a 250V process, named Multipower BCD, integrating Bipolar, CMOS and H.V. P-channel signal devices with DMOS power stages is demonstrated. The adopted integration scheme allows wide flexibility in the electrical output configuration. The process and device key features are presented.

1. INTRODUCTION

In Smart Power IC technology, to implement high voltage and power output functions, DMOS devices are preferred to Bipolar devices because of their simplified drive circuitry requirement, very fast switching speed, better reverse safe operating area (SOAR) and compatibility with CMOS logic.

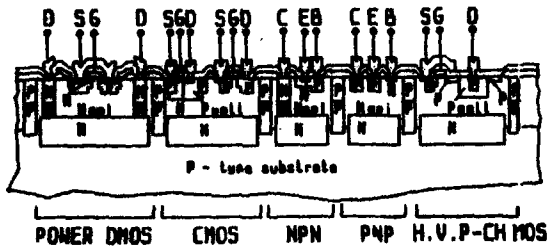
The integration of the DMOS transistor has been approached in various ways, depending on the final application and the required voltage and current capability.

Within a junction isolation technique, the lateral approach (RESURF type) is more suitable to achieve very high voltage capability but, in some cases, the source voltage swing is limited. The vertical design, with the drain contact on the back of the chip, is more suitable for high current requirement but it is limited to common drain integrated devices. A vertical structure with the drain contact brought to the top via N+ buried layer and sinker diffusions allows the integration of more than one high voltage medium current DMOS device, with no limit in the electrical output configuration. Following the last approach a technology termed Multipower BCD has been developed in the 60V [1], 100V and 250V ranges. The 250V rating, investigated in this paper, is suitable to develop off line motor controls and switched mode power supplies, solid state relays and intelligent switches, lamp ballasts and fluorescent display drivers. A test pattern has been designed to evaluate the process feasibility and the device performances. The design has been supported by the 2-D device simulator HFIELDS.

2. PROCESS KEY FEATURES

The architecture of the Multipower BCD process (fig.1) is based on the merging of the DMOS silicon gate process with the junction isolation technique. In the 250V version the junction isolation technique is improved by the top-bottom approach to reduce the high temperature drive-in times required by the needed thick epitaxial layer minimizing the buried layer outdiffusion and the isolation and sinker well sizes.

A large group of basic devices such as Bipolar, CMOS, DMOS, H.V. P-channel MOS transistors and zener diodes can be



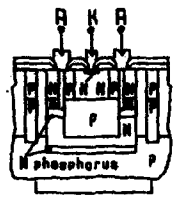


FIGURE 1

Multipower BCD 250V process cross section

obtained by using the dopant profiles involved in the DMOS silicon gate process adding only two extra P doped layers:

- the P well regions designed to realize self isolated N-channel MOS in CMOS devices, implanted after epitaxial growth and driven along with the top isolation and top sinker;
- the lightly boron doped drain extensions for the H.V. P-channel MOS transistors implanted before arseric doped sources, drains and emitters.

The total number of masking steps is 14 and a 4-um linewidth and 1-um alignement tolerance process is used.

3. DEVICE DESCRIPTION

3.1 Bipolar and CMOS devices

As regards the signal devices, whose main electrical characteristics are reported in tab.1, two kinds of NPN's are available according to the doping profile employed as a base region. For the lateral PNP's,

TABLE 1

SIGNAL DEVICES

BIPOLARS	h_{fe}	LV_{ceo} (V)	BV_{cbo} (V)	BV_{ebo} (V)
NPN1 P-well base	200	75	180	18
NPN2 P-body base	35	120	130	8
LPNP (L = 18u)	60	80	150	190

CMOS	V_{th} (V)	BV_{dss} (V)	g_m/Z (S/cm) Vgs=10 Vds= 5	$R_{on}\cdot Z$ ($\Omega\cdot$cm) Vgs=10 Vds=0.1	T_{delay} (nsec) f.o.=1
N-ch.	1.0	18	0.16	1.9	24 Vdd= 5
P-ch.	1.2	20	0.05	24	10 Vdd =10

spacings between emitter and collector wells down to 8-um are possible according to the gains and Early effect required.

A proper P well region for CMOS devices provides the matching between the N- and P-channel threshold voltages without any extra adjust implant. With the lithographic rules employed in the process, a density of about 450 tr/mm2, N- and P-channel mixed, can be obtained.

3.2 Low Leakage Diode (LLD)

If to the 14 mask process an extra mask is added to implant an optional phosphorus doped buried layer, the top-bottom approach allows the integration of an isolated high voltage LLD (fig.1) that is useful as a free-wheeling diode in inductive load driving.

This structure has achieved a blocking

voltage capability of about 290V and a current loss towards the substrate of about four orders of magnitude lower compared to a traditional base/collector diode, due to the very low gain of the parasitic substrate PNP.

3.3 N-channel Vertical DMOS transistor

The DMOS design work has been carried on to obtain a device that guarantees the desired voltage capability with the minimum RonxArea value, keeping into account some constraints. To avoid channel punch-through at the maximum applied voltage, the body region has a doping profile designed for a channel length of about 1.7-um and a threshold voltage of 3V. The epitaxial layer thickness is fixed equal to 25-um as a maximum value available by the top-bottom technique. The resistivity is chosen equal to 11-ohmxcm as a good tradeoff between the blocking requirement and the drift region contribute to the on resistance.

The junction edge termination that optimizes the device blocking voltage capability has resulted to be a triplanar field plate (fig.2) obtained extending the source metal beyond a biplanar polysilicon electrode running over the gate and the field oxide.

These two oxides, whose thicknesses are respectively 850-A and 2.5-um, are connected with a very low angle (less than 20°) to reduce the electric field peak arising at the transition point and to improve step coverage. The positioning of the third field plate level on an increased dielectric thickness lowers the induced electric field at the silicon surface, allowing a breakdown voltage 20V higher with respect to the biplanar one and making the device unaffected by field oxide thickness reduction from 2.5-um to 1.5-um.

However in the actual process, to reduce vertical and fringing electric fields induced by the high voltage interconnection crossover, a 2.5-um thick oxide is preferred.

The integrated DMOS layout configuration has been chosen according to a mathematical model [1] that provides the specific on resistance as a function of the physical and geometrical parameters involved in the structure (fig.3). Test vehicles designed

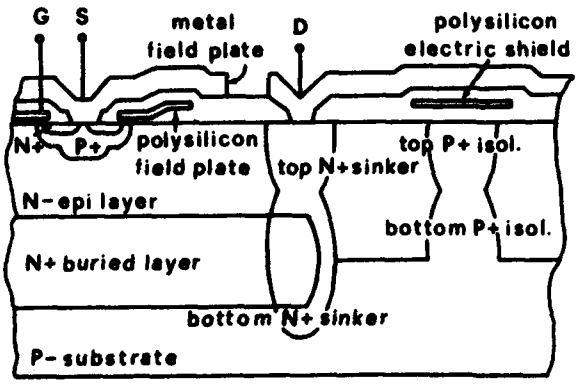


FIGURE 2

DMOS edge cross section

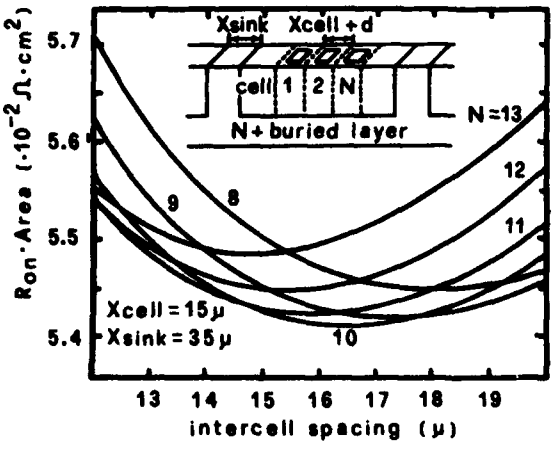


FIGURE 3

RonxArea vs. DMOS design parameters

placing between two drain sinker contacts 11 square shaped (15-umx15-um) cells with a gate width of 15-um have exhibited a specific on resistance equal to 5.2-ohmxmm2, in good agreement with the calculated value.

The DMOS electrical characteristics are shown in tab.2.

TABLE 2

HIGH VOLTAGE DEVICES

	V_{th} (V)	BV_{dss} (V)	g_m/Z (S/cm) $V_{gs}=10$ $V_{ds}=5$	$R_{on} \cdot Z$ ($\Omega \cdot$cm) $V_{gs}=10$ $V_{ds}=0.1$	$R_{on} \cdot Area$ ($\Omega \cdot cm^2$) $V_{gs}=10$ $V_{ds}=0.1$
P-ch. MOS	1.2	320	0.037	44	
N-ch. DMOS	3	315	0.11		$5.2 \cdot 10^{-2}$

3.4 High Voltage P-channel MOS transistor

The Multipower BCD process allows to integrate a H.V. P-channel MOS that can be used to drive the H.V. DMOS. The blocking capability is provided by an implanted extended drain that depletes back when a reverse voltage is applied avoiding electric field crowding and MOS channel punch-through.

The drain extension doping charge is gradually increased from the gate to the contact region, interposing the same medium doped P well profile employed in the CMOS, to better partition the applied voltage [2] .

Besides, the high surface electric field at the drain/gate overlap is reduced extending the source metal over a part of the more resistive drain extension region. A structure designed with drain extension and P well lengths respectively equal to 30-um and 10-um can sustain the required voltage on a wide range of drain extension implant doses (fig.4). Its electrical characteristics are reported in tab.2.

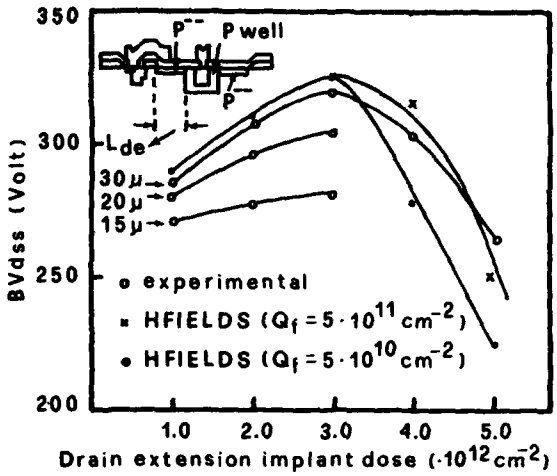


FIGURE 4

H.V. P-channel MOS BVdss vs. drain extension implant dose

4. CONCLUSIONS

By using a top-bottom junction isolation approach, low voltage Bipolars and CMOS, high voltage P-channel MOS and Low Leakage Diodes have been integrated with high voltage and medium current Vertical DMOS devices with planar contacts. The breakdown voltage of 300V for DMOS and P-channel MOS was respectively achieved by a triplanar field plate and a double doped drain extension as junction edge terminations.

REFERENCES

[1] Andreini, A. Contiero, C. Galbiati,P., IEEE Trans.on Electron Dev. ED-33 No.12 (1986) 2025

[2] Temple, V.A.K., IEEE Trans.on Electron Dev. ED-30 No.8 (1983) 954

RAPID THERMAL PROCESSING OF POLYSILICON EMITTER BIPOLAR TRANSISTORS IN A COMBINED CMOS/BIPOLAR PROCESS

L.A.Grant, D.W.McNeill, *P.F.Blomley

STC Technology, London Road, Harlow, CM17 9NA, UK
*LSI Logic Ltd., Grenville Pl., The Ring, Bracknell, Berks., RG12-1BP, UK

The effect of RTA time on the emitter profiles and base current of polysilicon emitter bipolar transistors has been studied. Experimental results show increasing base current with anneal time. The contact saturation current density Jos has been extracted for a device with polysilicon doping level of $3x10^{20}$ cm^{-3} and a 45 second 1100°C RTA.

1. INTRODUCTION

Polysilicon emitter bipolar transistors have been fabricated as part of a combined CMOS/Bipolar process described previously [1] [2]. Rapid Thermal Processing (RTP) serves the combined purpose of a) arsenic activation and drive-in for shallow emitter formation, b) anneal/activation for source and drain areas and c) flow and reflow of BPSG for planarisation and contact hole rounding.

Several recent publications [3][4] have analysed devices in which the emitter has been formed by arsenic out-diffusion from polysilicon. These report the effect of conventional furnace annealing conditions. In this paper the effects of high temperature processing in a Heat Pulse 410T Rapid Annealing System are assessed.

2. EXPERIMENTAL PROCEDURE

2.1. Device Processing

An n-well was diffused into a p-type substrate to act as both the collector of the bipolar transistors and the isolation well for the p-channel MOSFETs. The bipolar devices' base dopant was implanted and then driven-in during the CMOS gate oxidation. Windows were defined in this oxide to form the emitter regions. The wafers were cleaned and, immediately following an HF dip etch, a layer of LPCVD polysilicon was deposited. The polysilicon was doped by arsenic implantation at 40keV. Subsequent thermal processing consisted of a 45 minute 900°C furnace anneal/oxidation and two Rapid Thermal Anneals (RTA). Finally, after metal deposition and patterning, the wafers were annealed in N_2:H_2 at 430°C for 15 minutes.

In experiments the total RTA time was varied between 20 and 45 seconds. Also, on samples given a 45 second RTA, two emitter doses were investigated.

2.2. Electrical Measurements

Probing of a device with emitter dimensions of 65µm x 7.5µm was carried out on a thermally stabilised chuck at 300°K. Saturation currents Ibo and Ico were obtained by extrapolating from the measured Gummel plots, using an ideal diode gradient, figure 1.

2.3. Analysis

Secondary Ion Mass Spectrometry (SIMS) analysis was carried out on test structures which were processed along with the devices. Spreading Resistance Analysis (SRA) was also used to check the absolute dopant concentrations given by SIMS analysis.

3. MODELLING

3.1. Process Modelling

Simulation using SUPREM3 [5] was employed to predict the dopant profiles. Using default values, arsenic out-diffusion from the polysilicon into the single crystal was found to be in poor agreement with the measured profiles. Modelling the polysilicon as a constant gas source of arsenic in nitrogen, however, produced better results. This approach can be justified when one considers that arsenic diffuses very rapidly along the grain boundaries and establishes a constant average dopant concentration in the polysilicon. Furthermore the amount of arsenic in the single crystal part of the emitter is small amounting to less than 5% of the total dose in the polysilicon.

3.2. Device Modelling

Characterisation of polysilicon contacts to bipolar transistors has usually been carried out by use of the parameter Sp [6], the effective recombination velocity of the minority carriers at the poly/single crystal interface. For this work, however, it was decided to adopt the approach laid out in [4], and use the parameter Jos, the contact saturation current density as the figure of merit for the polysilicon contact. Jos is defined as

$$Jos = qS_p p_o(W_E)$$

where q is the electronic charge and $p_o(x)$ is the equilibrium hole concentration. The interface ($x=W_E$) is at a distance W_E from the emitter base junction ($x=0$).

The emitter saturation current density Joe, was taken as

$$Joe = Ibo/A_E$$

where A_E is the emitter area. The device simulator BIPOLE [7] was then used to extract Jos for the given doping profile by fitting the simulation result for Joe to the measured value by altering Sp. During this exercise the mobility data of Roulston et al [8] was used with an Auger recombination coefficient of $1.8\times10^{-31}cm^6s^{-1}$ employed on the basis of the work by del Alamo et al [9].

4. EXPERIMENTAL RESULTS

SIMS profile results for four different RTA times are given in figure 2. Arsenic is known to segregate to grain boundaries in the polysilicon. Therefore, a peak in the profile exists at the grain boundary between the poly and single crystal silicon. The effect of increased RTA time leading to increased arsenic out-diffusion can clearly be seen.

Figure 3 shows the combined SIMS, SRA and SUPREM3 profiles in both the emitter and base regions of the device. Good agreement is seen in the single crystal part of the emitter and at the emitter base junction. Deviation of the SRA results from the SIMS near the base collector junction is likely to be caused by field crowding near to the surface of the bevelled sample. Due to grain boundaries and hence increased resistivity in the polysilicon an erroneously low concentration is given by SRA in this layer.

The emitter saturation current densities are plotted as a function of RTA time in figure 4. Joe can be seen to increase with increasing time at 1100°C. The sample with the increased arsenic concentration of $3.75\times10^{20}cm^{-3}$ exhibits a marginally higher Joe than the sample with $3.0\times10^{20}cm^{-3}$. This difference is however not greater than experimental error.

Extraction of Jos was carried out for the $3.0\times10^{20}cm^{-3}$ doped emitter which had a 45 second RTA. This gave a value of

$$Jos = (1.5 \pm 0.3)\times10^{-12}\ Acm^{-2}$$

with 70-80% of the injected hole current reaching the original poly/single crystal interface.

These results agree with those of Patton et al [4], in demonstrating the increase in Joe as emitter drive-in time is increased. Again referring to [4] the Jos value given here would correspond to an interface which has had the

original native oxide broken up and has substantially realigned. The work by Wolstenholme et al [3] would suggest that the initial 45 minutes at 900°C does not significantly break up the native oxide. This means that the 1100°C anneal is necessary for this interface break up.

5. CONCLUSION

In this work results have been presented on the effects of RTA time on emitter profile and saturation current density for polysilicon contacted bipolar transistors. Increasing times lead to higher Joe values and larger base currents. This may not be undesirable however, if increased reproducibility is obtained through reduction of the interfacial effects, while maintaining shallow junctions.

REFERENCES

[1] Baker, R.B., Hunt, R.G., McNeill, D.W., Blomley ,P.F., Scovell ,P.D., ESSDERC 1986.

[2] Baker, R.B., Hunt, R.G., McNeill, D.W., Blomley ,P.F., Scovell ,P.D., Inst. Phys. Electroics Group, Feb. 1987.

[3] Wolstenholme, G.R., Jorgensen, N., Ashburn, P. Brooker, G.R., J. Appl. Phys. 61(1), Jan. 87

[4] Patton, G.L., Bravman, J.C., Plummer, J.D., IEEE, ED-33, No. 11, Nov. 1986

[5] Ho, C.P., Hansen, S.E., Stanford Elec. Lab. Tech. Report No. 83-001, July 1983

[6] Benna, B., Meister, T., Schaber, H. Wieder, A.W., IEDM 1985 pp. 302-305.

[7] Roulston, D.J., Chamberlain, S.V., Sedgal, J., IEEE Trans. ED-19, p.809, 1972

[8] Arora, N.D., Hauser, J.R., Roulston, D.J., IEEE Trans. ED-29, pp.292-295, Feb. 1982

[9] del Alamo, J., Swirhun, S., Swanson, R.M., IEDM 85, pp.290-293.

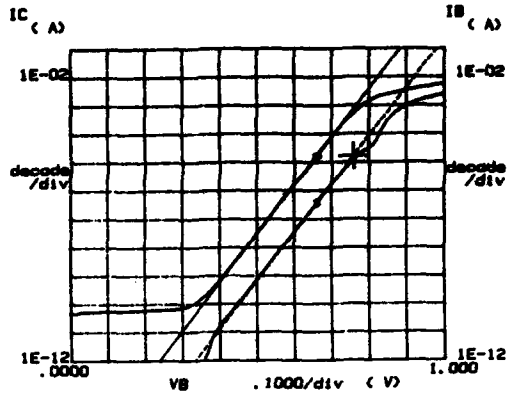


Figure 1

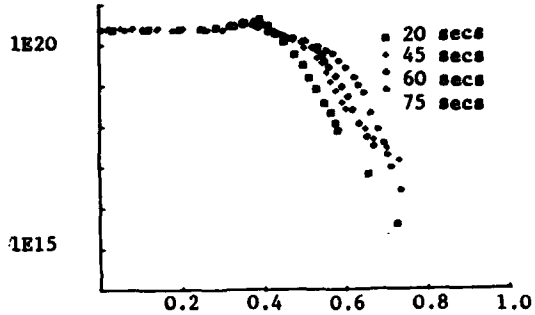


Figure 2

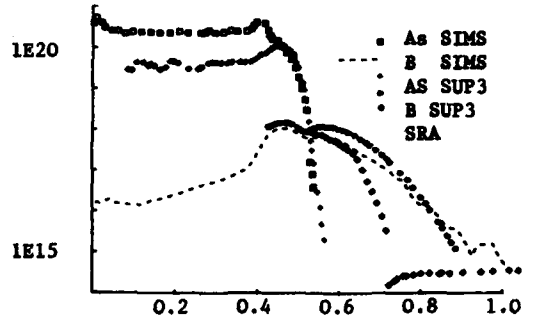


Figure 3

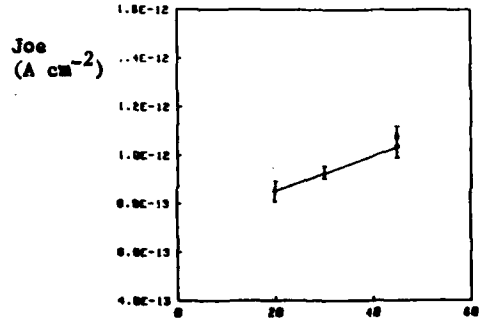


Fig. 4 RTA Time Secs

Session B1.1

Compound Semiconductors Technology I

Chairman: H. Martinot

Monday, September 14, 1987

SUBSTRATE INFLUENCES ON THE ACTIVATION OF ION-IMPLANTED Si IN GaAs

R.D. Schnell and H. Schink

Siemens AG, Corporate Research and Development
D-8000 München 83, P.O.Box 830952, F.R.G.

The influence of substrate material on Si-implanted CV-profiles is demonstrated on s.i. LEC grown GaAs. Local activation is affected by stoichiometry variations near dislocations. For higher concentrations the variations increase due to saturation effects. By co-implantation of As and enhanced implantation damage the activation is decreased and the dislocation influence increased.

Introduction

Silicon is the most commonly used dopant for the formation of n-type active layers by ion implantation into semi-insulating GaAs. The application of MESFETs for digital LSI circuits requires a high degree of threshold voltage uniformity and reproducibility. Therefore the doping profile, i.e. its shape and activation efficiency, has to be controlled carefully. However, the activation of the implanted silicon is dependent on substrate properties and quality. Dislocations, deviations from stoichiometry and concentrations of both gallium and arsenic vacancy defects as well as of impurities influence the activated profiles [1-5].

In the following paper we will demonstrate the influence of the substrate material itself on the doping profile and show the effect of an As^+ co-implantation. Macroscopic as well as microscopic variations of profiles in conventional undoped GaAs-substrates are evaluated quantitatively and interpreted in terms of their relevance for FET homogeneity and compared to requirements of LSI digital circuits.

Experimental

The substrate material used was Siemens 2"-semiinsulating undoped GaAs. Ion implantation was performed into the blanket wafer under appropriate tilt and rotation angles in order to avoid inhomogeneities due to planar channeling. Capless annealing was performed under an AsH_3-overpressure. A CV-pattern [6] (Fig. 1) was created by lift off technique using Ti/Pt/Au metallization. The CV-pattern allows for a high lateral resolution (100µm) because each Schottky contact is used nine times, once acting as the diode under test and eight times acting as part of the ring electrode. The interconnections are installed on the probecard. Measurements are performed on semi-automatic probestations using an LF-impedance analyser and applying appropriate corrections [6] for the evaluation of the CV-profiles. For mappings of the entire wafer every 10^{th} diode is measured, resulting in a resolution of 1mm.

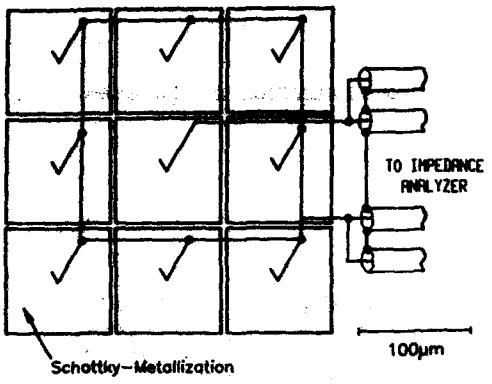


Figure 1

Schematic of the Schottky-metallization for high lateral resolution CV-measurements.

Results

Fig. 2 shows a capacitance map measured at zero bias voltage on a 2"-wafer implanted with ^{29}Si at 60keV with a dose of 3E12cm^{-2}. The material influence results in a higher capacitance in the center and at the edge of the wafer. The lowest values are located at the well known positions of lowest dislocation density along the [110] radial directions resulting in a four-fold symmetry. The variations are ±1.5% for a mean value of 1.37fFµm^{-2}. A comparison with the doping profile shows that these variations are due to variations in the overall activation efficiency. Since the width of the space charge region (W) is proportional to N$^{-1/2}$ (N=doping concentration), the variation of doping concentration $\Delta N/N$ is twice the variation of capacitance $\Delta C/C$. For the case mentioned above a variation in N of ±3% results. This is in agreement with the CV-profile. As the shape of the profile is constant, this means that V_T varies by the same amount, leading to values of ±18mV for V_T=+150mV. This is a value tolerable for LSI circuit applications [7].

On the microscopic scale the influence of dislocation networks on V_T was demonstrated frequently [1,3-5]. But there is little information about the doping profiles involved [6]. Especially the question whether an activation [1] or a compensation effect [2] is responsible is not clearly

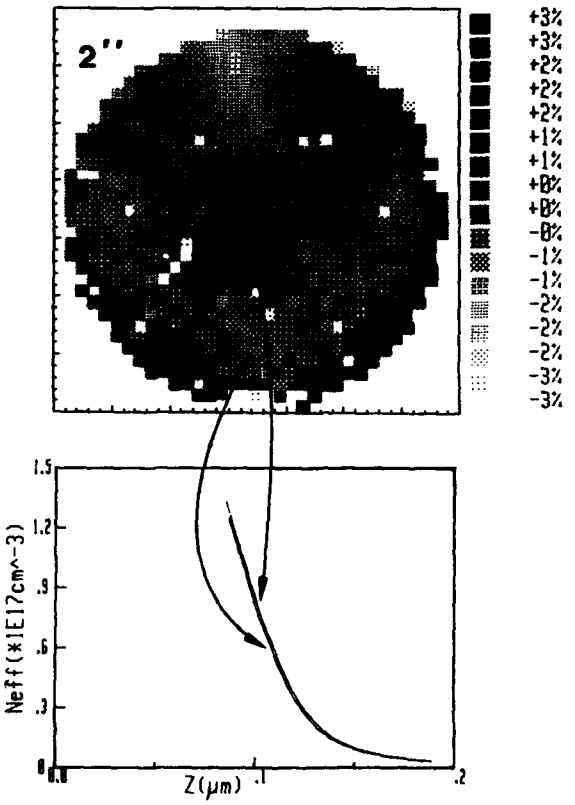


Figure 2

Capacitance map at 0V bias voltage (top) and selected CV-profiles (bottom) after Si implantation (60keV, 3*10^{12} cm^{-2}) into 2"-s.i. LEC grown undoped GaAs.

resolved. Fig. 3 shows a high resolution map of the capacitance of an area of 2x2mm² in the center of a wafer implanted at 60keV with a dose of 5E12cm⁻². The influence of the dislocation network is seen. The doping profiles taken at the center and the wall of one dislocation network differ from each other in such a way that the peak activation is higher at the wall than in the center. In the tail of the profile no differences are measured. This effect is enhanced by using higher implantation doses (Fig. 4). Going from

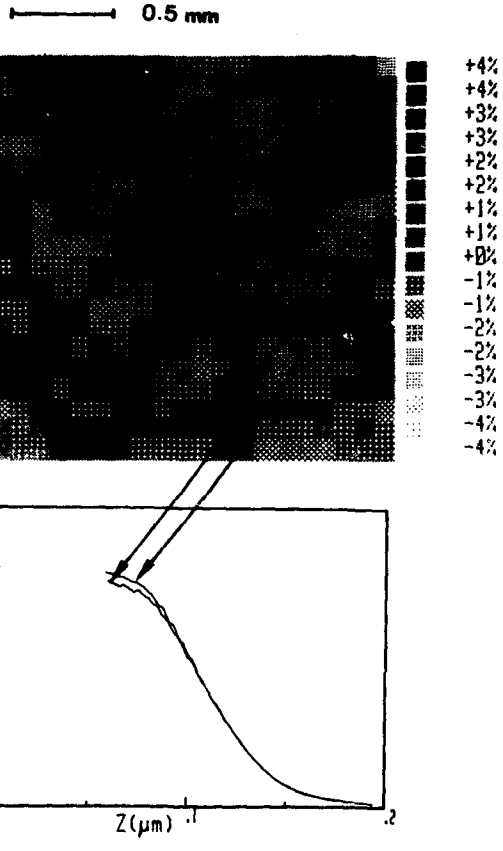


Figure 3

Microscopic capacitance map (top) and selected CV-profiles (Si^+,60keV,$5*10^{12}$ cm^{-2}, bottom) showing the influence of dislocation networks.

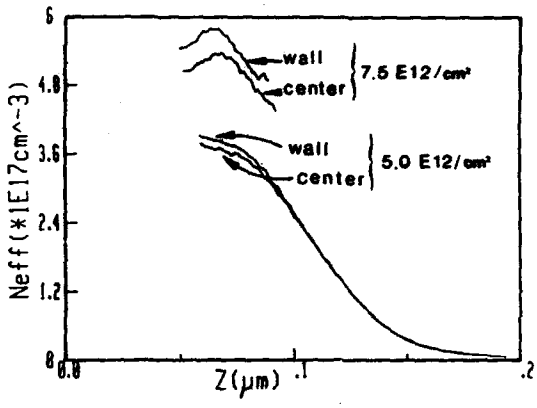


Figure 4

Dose-dependent microscopic activation scattering for 60keV Si implants.

5E12cm⁻² to 7.5E12cm⁻² the relative difference of the peak carrier concentrations increases from 3% to 7% (corresponding to V_T-values of -1.9V ± 20mV and -3.2V ± 70mV respectively). The results can be explained by a saturation of the Ga vacancy occupation with Si atoms that takes place at a lower Si concentration in the center of the network than at the wall. By approaching a saturation concentration of 1...2E18cm⁻², where the amphoteric character of Si prevents higher concentrations to be activated, the dislocation influence is enhanced.

A similar increase is observed by the use of an As^+ co-implantation of 1E13cm⁻² at 170keV into a Si-layer formed at 60keV with a dose of 7.5E12cm⁻² (see Fig.5). Both implantations where capless annealed together. In this case the activation efficiency decreased as a consequence of the implantation damage. The ratio $\Delta C/C$ between center and wall of the dislocation network was 6%.

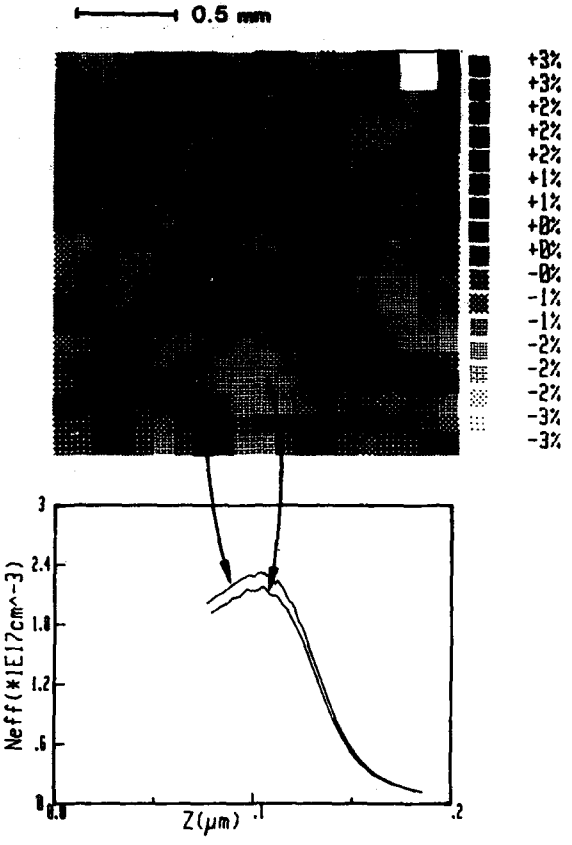


Figure 5

Microscopic capacitance map (top) and selected CV-profiles (bottom) after co-implantation of Si (60kV,$7.5*10^{12}$ cm^{-2}) and As (170kV,$1*10^{13}$ cm^{-2}).

Conclusions

The results can be interpreted in terms of the simple defect model of Miyazawa et.al. [1]. It was found that the concentration of EL2 [8] is increased around dislocations with respect to its average value. As a consequence the concentration of Ga-vacancies and the activation efficiency for Si increases. When the density of Si is increased the activation efficiency and the scatter of the activation become concentration dependent due to saturation effects. The lowest variations are obtained with an annealing process that results in nearly 100% activation as shown in Fig.3.

References

[1] S.Miyazawa and K.Wada, Appl.Phys. Lett. 48 (1986) 905

[2] S.Yasuami,K.Fukuta,M.Watanabe and T.Nakanisi, Jap.J.Appl.Phys. 25 (1986) 1905

[3] H.Schink,G.Packeiser,J.Maluenda and G.M.Martin, Jap.J.Appl.Phys. 25 (1986) L369

[4] G.Packeiser,H.Schink and H.Kniepkamp, Proc. Semiinsulating III-V Materials, Hakone, Ohmsha Ltd. (1986) p.561

[5] T.Egawa,Y.Sano,H.Nakamura and K.Kaminishi, Jap.J.Appl.Phys 25 (1986) L973

[6] H.Schink, Proc. E-MRS, Straßbourg (1987), to be published

[7] G.Packeiser, Proc. NATO-Workshop on Microscopic Inhomogeneities of Bulk GaAs, Oxford 1987

[8] H.C.Alt and G.Packeiser, J.Appl. Phys. 60 (1986) 2954

COMPARISON OF RAPID ANNEALING AND FURNACE ANNEALING OF SI IMPLANTED INTO GAINAS.

J. Splettstößer, H. Heesel, U. Breuer, W. Albrecht ,
D. Schmitz*, J. Selders** and H. Beneking***

Institute of Semiconductor Electronics
Aachen Technical University
Sommerfeldstraße, D-5100 Aachen, FRG

The activation efficiency of conventionell furnace annealing (650 °C-900 °C, 30 min, SiO_2 cap, N_2 ambient) and rapid thermal annealing (600 °C-900 °C, with and without SiO_2 cap, N_2 ambient) has been compared in Si doped GaInAs layers grown lattice matched by either OMVPE or LPE on InP. Carrier profiles and atomic profiles have been determined by selective Hall measurement, C/V profiling and SIMS measurements. For an implantation dose of $2*10^{14}$ cm^{-2} and rapid thermal annealing at 900 °C activation of 69 % with a sheet resistance of 20 Ω is found. Rapid thermal annealing and furnace annealing lead to comparable activation efficiency and Hall mobility data. A broadening of carrier concentration profile for furnace annealing at 700 °C, 30 min is observed. Rapid thermal annealing between 700 °C and 900 °C leads to no measurable change in Si profile.

In case of low dose Si implantation in OMVPE grown layers highest activation (65 %) is achieved for rapid thermal annealing at 700 °C. The electron mobility at a doping concentration of 10^{17} cm^{-3} is 5800 $cm^2v^{-1}s^{-1}$. For low dose Si implantation in LPE grown layers the activation is 100 % after rapid annealing at 800 °C. Long furnace annealing (700 °C, 30 min) results in an anomalous high carrier concentration at the surface.

1. INTRODUCTION

Si can be used to produce highly doped n-GaInAs layers as desired for JFET and MISFET source and drain regions[1]. However the minimization of diffusion of dopants into GaInAs bulk and along the insulator/GaInAs interface requires a reduction of annealing temperature and annealing time. Furthermore we want to explore the possibilities of Si ion implantation for moderately doped layers ($n=10^{17}cm^{-3}$) as required for channel regions [2].

The aim of our work is to optimize annealing conditions with regard to activation and recrystallization in order to avoid diffusion.

2. EXPERIMENTAL

We studied the annealing of Si implanted into n- and p-type GaInAs with doses between $1*10^{12}$ cm^{-2} and $3*10^{14}$ cm^{-2}. The implantation energies varied between 25 keV and 200 keV. Lowly doped GaInAs has been grown by LPE and OMVPE lattice matched on InP substrates. The distribution of implanted Si has been measured by SIMS, selective Hall measurement and C/V profiling. In case of low dose Si implantation 2 K PL has been used to monitor the optical quality of the layers. The implanted layers have been annealed either by rapid thermal annealing or furnace annealing. The temperature was varied in the

* Aixtron, D-5100 Aachen, FRG
** Present adress: Telefunken Electronic GmbH, D-7100 Heilbronn, FRG
*** University of Michigan, Department of Electrical Engineering and Computer Science, Center for Highfrequency Microelectronics, 1301 Beal Avenue, Ann Arbor, Mich. 48109-2122/USA

range from 600 °C to 1000 °C. Rapid thermal annealing was performed with a tungsten lamp furnace. The heating rate is about 100 °C per second. Anneal time is defined as the period during which temperature is higher than 90 % of anneal temperature. All rapid thermal annealings were performed with the shortest possible anneal time, which is 3 s . In case of RTA the implanted samples were either encapsulated with SiO_2 or capless conditions were chosen.

3. RESULTS AND DISCUSSION

In Fig. 1 the influences of annealing temperatures (methods RTA or 30 min furnace annealing) and capping conditions (capless or with a SiO_2) are plotted.

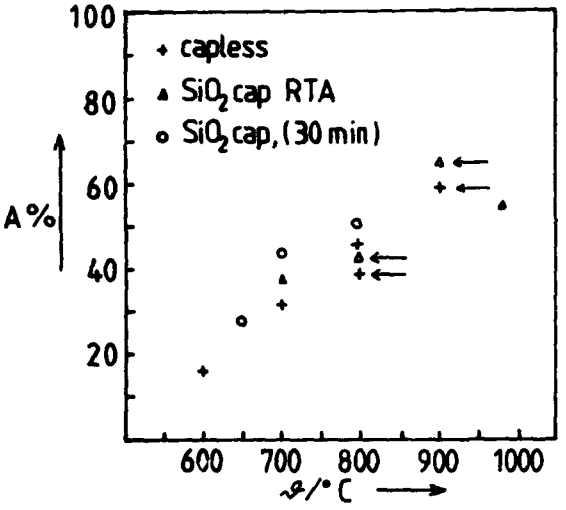


FIGURE 1

Dependence of electrical activation for high dose Si implanted GaInAs layers on annealing temperature. The implantation is a double implantation E=100 keV, D=6*10^{13} cm^{-2} and E=50 keV, D=3*10^{13} cm^{-2}. Capless and annealing with SiO_2 cap as well as RTA and furnace annealing are compared. Results obtained on OMVPE grown GaInAs layers are marked with arrows. In this case the implantation parameters are E=100 keV and D=2*10^{14} cm^{-2}.

As usually observed the activation increases with higher annealing temperatures [3] . At the same temperature furnace annealing results in a 10 % higher activation than rapid thermal annealing. The influence of cap on activation is low. The activation in LPE grown samples is 10 % higher than in OMVPE grown material. In Fig.2 carrier concentration profiles received from selective Hall measurements are compared with the LSS model.

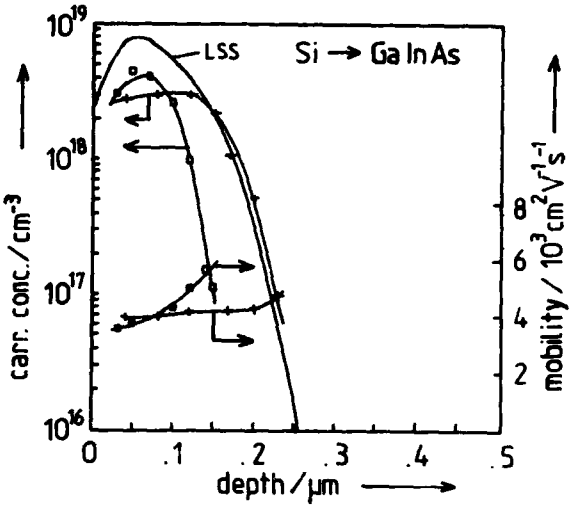


FIGURE 2

Electron concentration and electron mobility profile measured by selective Hall measurement. The profile resulting from furnace annealing (700 °C, 30 min)is marked with crosses. Squares represent the profile from rapid annealing at 700 °C.

The profile resulting from RTA at 700 °C suggests that diffusion is negligable. However furnace annealing broadens the profile.

SIMS measurements are used to determine projected range and projected standard deviation values for Si implanted with energies between 25 keV and 200 keV. For sputtering 10 keV Cs^+ ions or 10 keV O_2^+ ions have been used.

The experimental range data have been determined by a Gauss fit and are listed in Table 1 in comparison with data from LSS model. Especially the high energy implants correspond with theory.

TABLE 1

Range data for Si implantation with different energies.

		LSS theorie	experim.
50keV	$R_p/\mu m$	0.044	0.037
	$\Delta R_p/\mu m$	0.028	0.033
100keV	$R_p/\mu m$	0.087	0.075
	$\Delta R_p/\mu m$	0.048	0.050
200keV	$R_p/\mu m$	0.177	0.184
	$\Delta R_p/\mu m$	0.082	0.108

SIMS depth profiles of a Si implantation (E= 50 KeV; D= $2*10^{15}cm^{-2}$) are shown in Figure 3. Rapid annealing between 700 °C and 900 °C introduces no measurable change in Si profile.

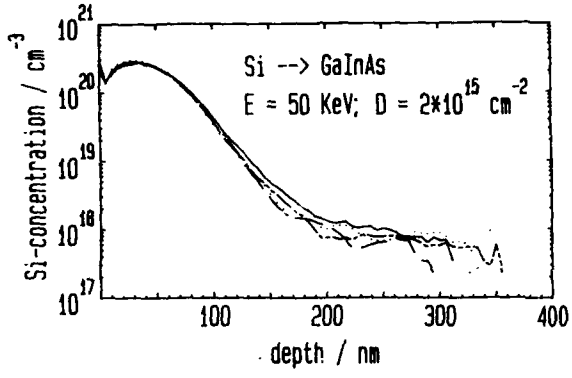


FIGURE 3

Si depth profiles for rapid thermal annealed samples measured by SIMS (——— as impl., ····· 700°C, — -- 800°C, — — 900°C).

Carrier concentration profile and mobility profile for low dose implantation in OMVPE grown samples are shown in Figure 4 and 5. It is remarkable that the mobility at a electron concentration of $10^{17}cm^{-3}$ is as high as 5800 $cm^2v^{-1}s^{-1}$. In case of rapid thermal annealing at 800 °C and 900 °C the activation is reduced to 45% and 36%, respectively. The decrease o activation with increasing anneal temperature is only observed for the low dose implantation in OMVPE grown layers.

The carrier concentration profile caused by long time furnace annealing (700 °C, 30 min., SiO_2 cap) is shown in Figure 6. An anomalous high carrier concentration is obtained for the

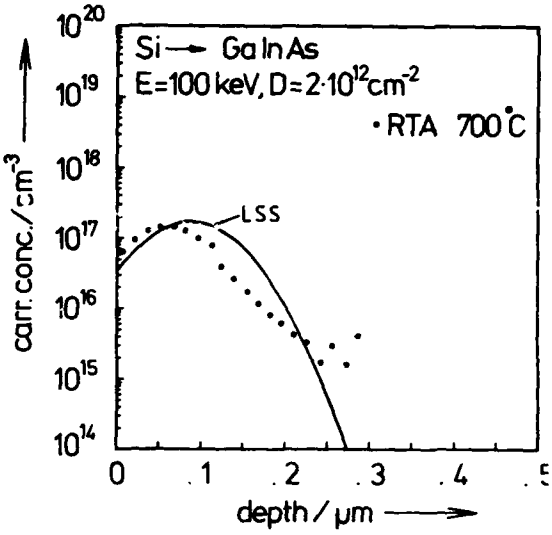


FIGURE 4

Carrier concentration profile of a low dose Si implanted sample, obtained by selective Hall measurement.

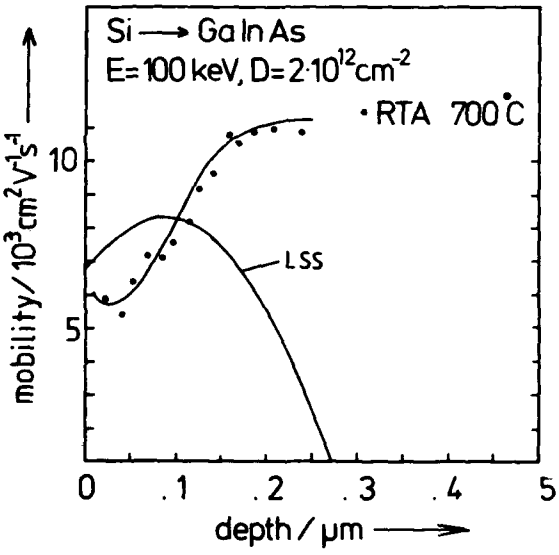


FIGURE 5

Mobility profile corresponding to figure 4.

surface region. This can be caused by Si indiffusion or stress induced donor like defects due to annealing with SiO_2 cap [4].

Optical quality of the low do . Si implantation has be n investigated by PL. Typical 2 K PL spectra of LPE grown Si implanted GaInAs layer excited by the 647,1 nm line of a Krypton

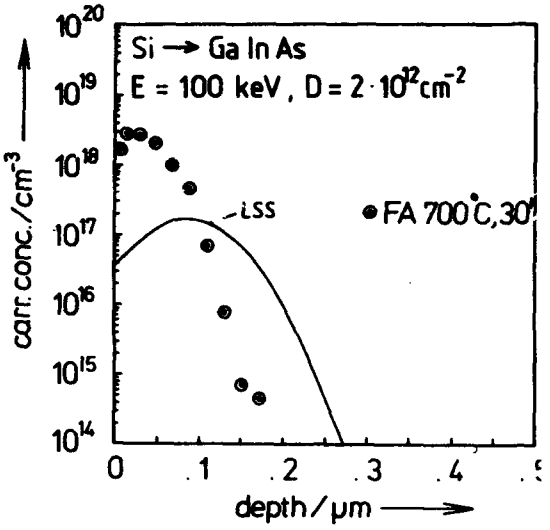


FIGURE 6

Carrier concentration profile resulting from a long time furnace annealing (700 °C, 30 min).

laser are shown in Figure 7. The excitation level for all curves is 0.6 W/cm^{-2}. Line A is correlated to the exciton bound to a neutral donator (D^o,X) and to the free exciton emission (X). Line B is attributed to a donator-acceptor (D,A) transition[5]. After rapid annealing at 700 °C the integral PL yield is comparable with the yield from the as grown sample. Higher annealing temperatures lead to a further increase of the PL intensity.

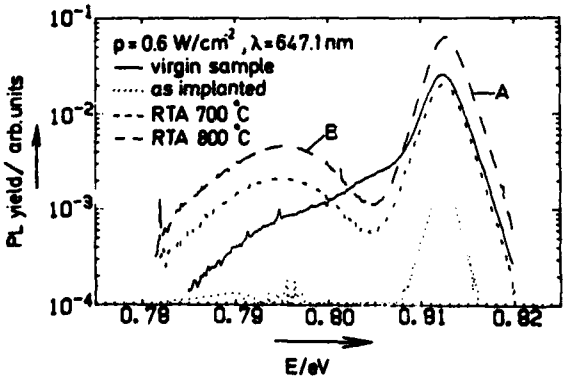


FIGURE 7

PL spectra of a as grown sample and an implanted sample before and after rapid annealing.

4. CONCLUSION

Doping levels of $10^{19}cm^{-3}$ can be achieved by Si implantation and rapid thermal annealing. For an implantation dose of $2*10^{14}cm^{-2}$ and rapid thermal annealing at 900 °C activation of 69 % with a sheet resistance of 20 Ω is found. The activation of high dose implantation is about 10 % lower in OMVPE grown layers than in LPE grown samples. Rapid thermal annealing and furnace annealing lead to comparable activation efficiency and Hall mobility data. A broadening of carrier concentration profile during furnace annealing at 700 °C, 30 min. is observed. Rapid thermal annealing between 700 °C and 900 °C leads to no measurable change in Si profile.

In case of low dose Si implantation in OMVPE grown layers highest activation (65 %) isachieved for rapid thermal annealing at 700 °C. It is remarkable that the mobility at a doping concentration of 10^{17} cm^{-3} is as high as 5800 $cm^2V^{-1}s^{-1}$. For low dose Si implantation in LPE grown layers the activation is 100 % after rapid annealing at 800°C.

ACKNOWLEDGEMENTS

The authors wish to thank S. Ambros for PL measurements and discussion andD. Dunkmann for the ion implantation.

This work has been supported by the EEC.

REFERENCES

[1] J. Selders and H. Beneking, IEEE Electron Device Lett., vol. EDL-7, p. 434, 1986.

[2] J. Selders, P. Roentgen and H. Beneking, Electronics Letters, vol.22, p. 14, 1986.

[3] T. Penna, B. Tell, A.S.H. Liao, T.J. Bridges and G. Burkhardt, J. Appl. Phys., vol. 57, p. 351, 1985.

[4] K.S. Seo, P.R. Berger, G.P. Kothiyal and P.K. Bhattacharya, IEEE Trans. Elec. Dev., vol. ED-34, p. 235, 1987.

[5] K.-H. Goetz, D. Bimberg, H. Jürgensen, J. Selders, A.V. Solomonov, G.F. Glinski and M. Razeghi, J. Appl. Phys., vol. 54, p. 4543, 1983.

YIELD-PERFORMANCE CONSIDERATIONS FOR ION-IMPLANTED GaAs INTEGRATED CIRCUITS BASED ON SUBSTRATE MATERIAL PROPERTIES.

C.Lanzieri, R.Graffitti, C.Calori, S.Rapisarda, and A.Cetronio

SELENIA Industrie Elettroniche Associate, Direzione Ricerche, Via Tiburtina, 00131 Roma, Italy.

ABSTRACT

In this article we will illustrate, in terms of yield-performance considerations, how for conventional annealing techniques the doped LEC materials (i.e. lightly Cr or In-doped) prove to be better than the undoped material, primarly because of the higher yield capability, and that before full advantage can be taken from the potentially better undoped material then either an improved annealing technique and/or ingot annealed material must be considered.

INTRODUCTION

GaAs integrated circuits are attracting much attention due to their high speed and high frequency capabilities. Recent progress in these IC's has mainly been attributed to the advance of direct ion-implantation into semi-insulating bulk GaAs and the corresponding post-implant anneal. Althougth the ion implantation technique has a high potential for realising uniform and reproducible active layers, said property is in part invalidated by the fact that not all GaAs substrates are equally uniform and furthermore they behave differently with post implant annealing and this often results in poor overall yields of IC's which do not respond to the specified performance limits. Because from the point of view of production yield the on wafer radial variations tend to present the biggest problems, in this work we have investigated how such variations can differ from one type of semi-insulating material to another on the basis of an evaluation which takes into account the devices specified electrical performance and yield.

In fact by means of electrical parameter mapping of arrays of field effect transistors fabricated on Liquid Encapsulated Czochralsky (LEC) grown semi-insulating GaAs, we illustrate, in reasonable agreement with other published work (1-3), how the on wafer active layer uniformity is strongly dependent of substrate material used, and furthermore how said uniformity can vary both radially on a wafer and longitudinally along the ingot from seed to tail-end. In particular we will illustrate, in terms of yield performance considerations, how for conventional annealing techniques the doped LEC materials (i.e. lightly Cr or In-doped) prove to be better than the undoped material, primarily because of the higher yield capability, and that before full advantage can be taken of the potentially better undoped material then either improved annealing techniques or ingot annealed material must be considered.

EXPERIMENTAL.

The material studied in this work includes five different types of S.I. GaAs substrates grown by the LEC technique: that is, undoped,

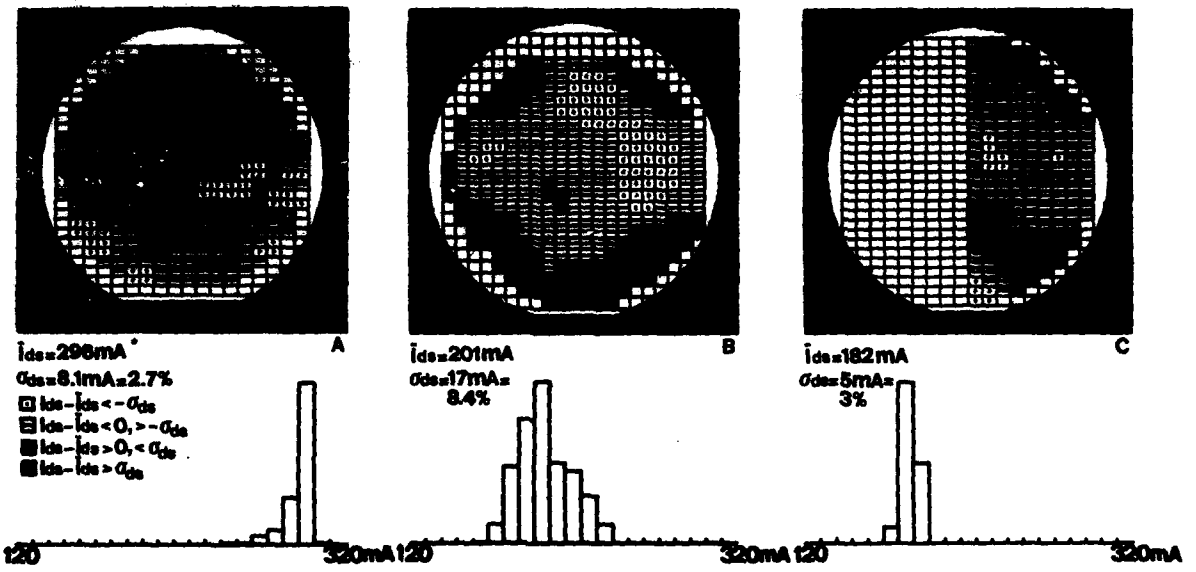


Fig.1 - On wafer activation uniformity maps and I_{DSS} histograms for different LEC substrates after conventional furnace anneal: a) Cr-doped, b) undoped, and c) undoped ingot annealed.
* Implant dose $1x10^{13}$ 60 keV + $6x10^{12}$ 150 keV.

undoped ingot annealed, Cr/O doped, In-doped and Cr/O doped with buffer layer. Said material was qualified by our standard "in-house" procedure to evaluate its properties (4). In particular the thermal stability of the material (i.e. resistance to surface conversion), after treatment with the "in-house" capping and annealing procedure and the activation efficiency and on wafer uniformity, after ion-implantation are evaluated.

Post-implant annealing of the substrates was carried out in the presence of a Silicon Nitride cap deposited by a reactive sputter deposition technique, and controlled to satisfy the "no-conversion" (5) criterion. The annealing cycles investigated include a standard furnace anneal at 820°C for 15 minutes in a nitrogen gas ambient and a modified technique which minimises the presence of thermal gradients across and through the GaAs substrates during annealing.

For this study we have used a standard $^{29}Si^{+}$ implantation to obtain the selectively doped active layers; in particular a double implant dose of $1x10^{13}$ at 40 keV and $5x10^{12}$ at 120 keV, typical for our medium to low-noise MESFET devices, has been utilised. Said implant dose has been prefered for this study because it is sufficiently low, (i.e. in the linear part of the activation versus implant dose characteristic) (6) to ensure reasonable sensitivity to on wafer activation uniformity variations, avoiding the problems of doping impurity saturation threshold typical for implant doses greater than $2x10^{13}cm^{-2}$.

The on wafer activation uniformity was studied by of arrays of MESFET ohmic contacts fabricated on the selectively doped S.I. GaAs substrates (2), with corresponding measurement of device source-drain saturation current I_{DSS} (300 μm wide devices) uniformily distributed on the 2 inch wafer. After this first characterisation we then proceded to complete the MESFET's by the conventional recessed gate technology and finally we evaluate their electrical performance by suitable d.c. and r.f. measurements.

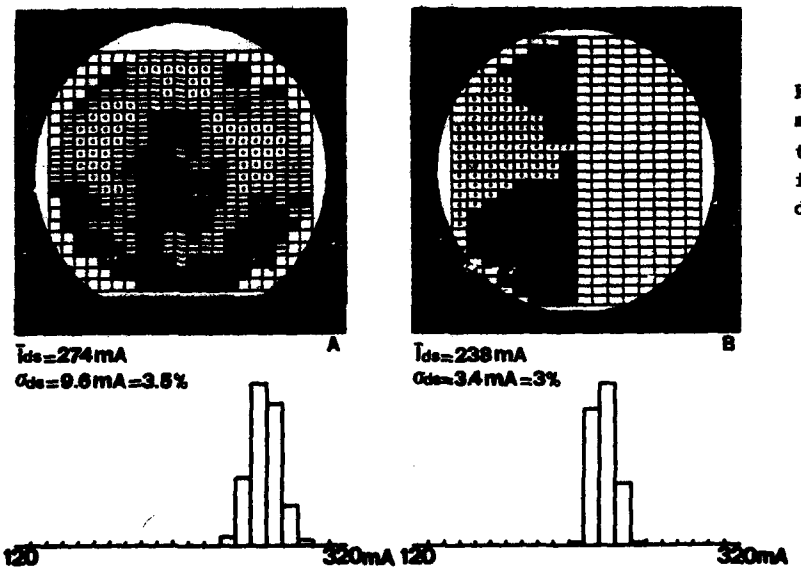


Fig.2 - On wafer activation uniformity maps and I_{DSS} histograms for the a) undoped, and b) undoped ingot annealed substrate after modified annealing conditions.

RESULTS.

From typical I_{DSS}, maps obtained for the different substrates after conventional furnace annealing, some interesting aspects of on wafer activation efficiency and uniformity are immediately apparent.

As shown in fig.1a for the doped LEC material (Cr doped in this case) the sheet carrier concentration is fairly uniform over the entire wafer with mean I_{DSS} standard deviation bettern than $\pm$ 3% irrespective of the wafer position in the ingot. However for this material the absolute mean I_{DSS} value (and thus activation efficiency) is found to decrease appreciable, by as 50%, in going from seed to tail end wafer.

As shown in fig.1b for the undoped LEC material a radial four-fold activation symmetry is obtained, similar to the maps of residual stress, dislocation and EL2 levels observed by near infrared transmittance (7), with a correspondingly poor on wafer activation uniformity. In fact for this material the mean I_{DSS} standard deviation was found to be in the best case (centre of ingot) $\pm$ 8% and in the worst case (tail end wafer) $\pm$17%. Furthermore for this material the absolute mean I_{DSS} value is unexpectedly found to decrease by as much as 30% in going from seed to tail end wafer.

For the undoped ingot annealed material, unexpectedly large differences in activation uniformity have been observed in going from seed to tail end wafer. In particular we have observed that even thougth the absolute mean I_{DSS} value is unchanged from seed to tail end wafer the corresponding standard deviation in mean I_{DSS} can be as poor as $\pm$ 30% for the seed end wafer and as good as 3% for the tail end wafer (shown in fig.1c). This apparently anomolous behaviour in activation uniformity of undoped ingot annealed seed-end wafer has been confirmed by similar results obtained from other ingots from different suppliers.

Finally for the Cr/O doped substrates with buffer layer we observe only localised disomogeneous areas, yielding typical mean I_{DSS} standard deviations better than $\pm$4%.

As might be expected the above reported results are strongly dependent on the post-implant

capping and annealing procedure. In fact as illustrated by some typical results presented in fig.2, we observe that by reducing as much as possible the thermal gradients generated across and througth the GaAs substrate during annealing (particularly during cooldown), the activation efficiency and in particular the on wafer uniformity of the undoped LEC material can be improved appreciable. In fact as shown in fig.2a the mean I_{DSS} standard deviation for the undoped material in this case is always better than $\pm$7%, with a constant activation efficiency throughtout the ingot, and as shown in fig.2b for the undoped ingot annealed material even thougth there is little improvement in the mean value standard deviation for the seed end wafer (remains $\pm$3%)the activation efficiency is found to increase by as much as 30%.

Preliminary r.f. measurements (i.e. optimum noise figure NF and maximum available gain G_{max} at 12 GHz) of the MESFET's fabricated with the above reported substrates tend to indicate that for conventional furnace annealing conditions there is little or no differehce in device performance with substrate material used. In fact both the Cr-doped and undoped material typically give NF = 2.3 dB with G_{ASS} 7 dB and G_{max} = 10dB. Obviously in this situation the Cr-doped material results in a better yield/performance evaluation as a result of its better activation uniformity. In fact the potential of the undoped LEC material, in terms of improved r.f. performance of the MESFET's was only evident with the improved annealing technique. For such a condition NF = 2.0dB and G_{max} = 11dB have been obtained, and this together with the improved activation uniformity for the undoped ingot annealed material (shown to be as good as $\pm$ 3% in fig.2b) results in good yield/performance evaluation for ion implanted MESFET devices.

CONCLUSION.

In our characterisation of S.I. substrates we have seen that by utilising a conventional furnace annealing technique, the yield/performance evaluation of MESFET devices fabricated on undoped LEC material is surprisingly inferior to that obtained for doped material, primarly because the undoped material gives a lower yield of devices within r.f. specifications. The surprising nature of this result has led us to investigate in more detail the post-implant annealing technique and as such have found that before full advantage can be taken from the potentially better undoped or undoped ingot annealed LEC material then an improved annealing technique, which minimises as much as possible the presence of thermal gradients across and througth the GaAs wafer during annealing, is necessary.

REFERENCES.

1. C.G.Kirkpatric, R.T.Chen, D.E.Holmes and K.R.Elliott; Gallium Arsenide: Material Devices and Circuits, Edt.by M.J.Howes and D.V. Morgan, J.Wiley & Sons Ltd. (1985).
2. P.Dobrilla, J.S.Blakmore, et al: Appl.Phys. Lett. 47(6), 602, (1985).
3. P.Dobrilla and J.S.Blakmore; J.Appl.Phys. 60(1), 169, (1986).
4. A.Cetronio, published on E.MRS Conference 1987 meeting "Symposium on Growth, Characterisation, Processing of III, V materials with correlations to Devices Performances", Edt. by P.A.Glasow, Y-I Nissin.
5. P.B.Klein, P.E.R.Nordquist and P.G.Seebenmann: J.Appl.Phys. 51(9), 4861 (1980).
6. S.G.Liu, E.C.Douglas, C.P.Wu, C.W.Magee, S.Y.Narayan, S.T.Jolly, F.Kolandra and S.Jain; R.C.A. Review, 41, 227 (1980).
7. P.Dobrilla and J.S.Blakmore; J.Appl.Phys., 61, 1442 (1987).

Session C1.1

Power Devices
I

Chairman: A. Stella

Monday, September 14, 1987

A NOVEL 1500 VOLT IGBT DEVICE WITH IMPROVED TURN-OFF PERFORMANCE

J. BLAKE, M.E. BROCKMAN, R.W. COOPER

Philips Research Laboratories, Redhill, Surrey U.K.

An Insulated Gate Bipolar Transistor design based on bulk silicon material capable of blocking voltages in excess of 1500V is reported. The device incorporates a detailed anode shorting pattern which inhibits hole injection during on-state, and reduces turn-off times.

1. INTRODUCTION

The Insulated Gate Bipolar Transistor (IGBT) combines the low on-resistance of a conductivity modulated device with the simple gate drive requirements of the power MOSFET. Several examples of this device have been reported [1-4]. These are based on $n:p^+$ epitaxial starting material with blocking voltages up to 1200V. Generally to reduce turn-off time and thereby enhance the switching speed some form of carrier lifetime control is required. In this work we report on IGBT devices capable of blocking voltages in excess of 1500V.

2. DEVICE DESIGN AND FABRICATION

The device design is shown schematically in Fig.1 with its equivalent circuit in Fig.2. The starting material is 100 Ω cm bulk silicon which is polished down to a thickness of approximately 250μm. This allows the blocking voltage to be attained without excessive forward voltage drop. The top processing follows conventional auto-aligned VDMOS schedules using a cellular mask design. However, this processing is preceded by back of slice implants which produce an array of p^+ dots in an n^+ back contact. The dot size and separation are chosen to be much smaller than the slice thickness allowing relatively coarse alignment of the top and bottom patterns.

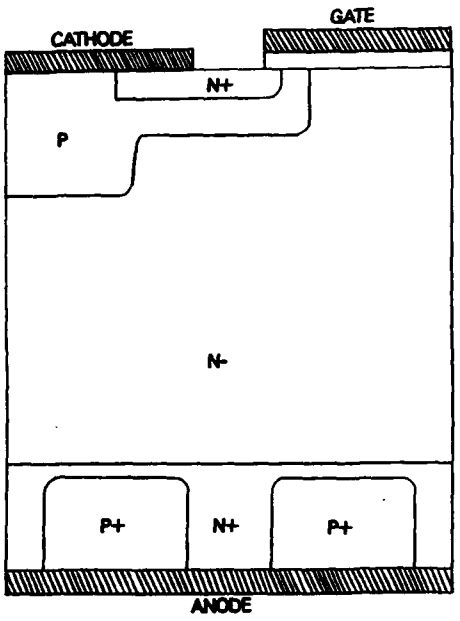


FIGURE 1
Schematic Device Structure

Different mask geometries lead to a variation in the back p^+ to n^+ area ratio.

This design then gives a shorted anode structure which serves to suppress injection during the on-state whilst additionally providing a path for electrons during the turn-off phase.

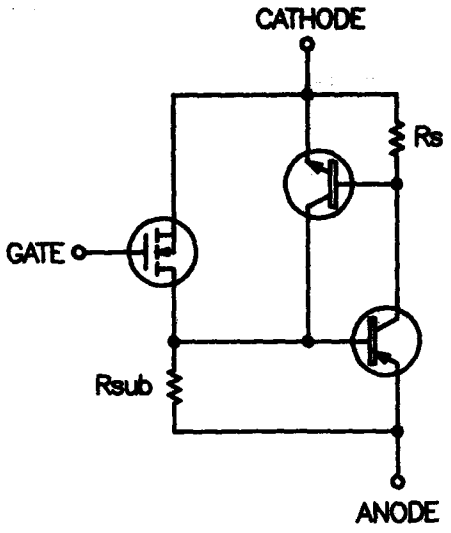


FIGURE 2
Equivalent Circuit

3. DEVICE MECHANISM

In the normal mode of operation a positive voltage is applied to the gate with respect to the cathode. When the anode voltage increases from zero we see conduction of electron current via the lateral n-channel FET and resistance R_{sub}. This differs from the usual I-V characteristic of an IGBT which shows no conduction until the anode voltage exceeds approximately 0.6V with respect to the cathode. Injection will occur from the back p^+ dots when the voltage across R_{sub} forward biases the emitter-base junction of the pnp transistor. Typically this will occur with a forward voltage of 1.2 to 0.8V for p^+ dot densities of 70 to 90%. Fig.3 shows the I-V characteristic of a device with p^+ density of approximately 85% showing injection at a voltage just above 0.8V. The n^- region now becomes modulated with holes from the p^+ anode dots, and electrons flowing through the inversion channel, lowering the on-resistance of the device. The resistance of the p-well, R_s, is minimised to short out the parasitic npn transistor thus avoiding device latch-up.

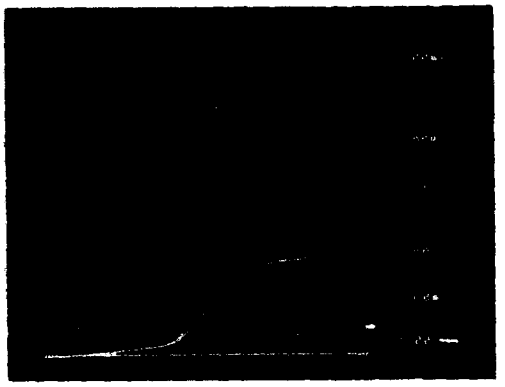


FIGURE 3
I-V Characteristic

4. TRANSIENT RESPONSE

Device turn-off is achieved by removing the positive gate voltage which shuts off the inversion channel. Now there is a direct extraction path for electrons at the anode and device turn-off is greatly enhanced. Current and voltage waveforms under resistive and inductive loads (Figures 4 and 5), show a storage time of approximately 200ns, then a rapid fall of about 50ns followed by a low level current tail of approximately 1.5µs.

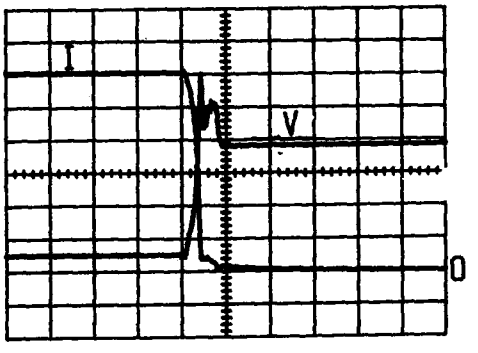


FIGURE 4
Anode Current and Voltage Waveforms with a Resistive Load
(0.5A/div, 5V/div, 500ns/div)

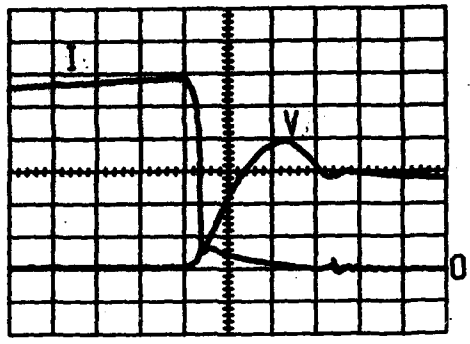


FIGURE 5

Anode Current and Voltage Waveforms with an Inductive Load
(0.5A/div, 50V/div, 500ns/div)

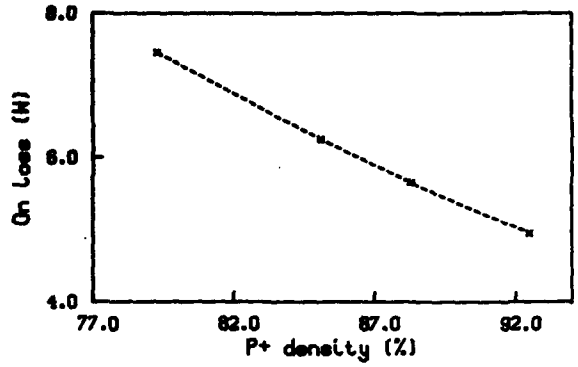


FIGURE 6

Variation in On-Losses with p^+ Density

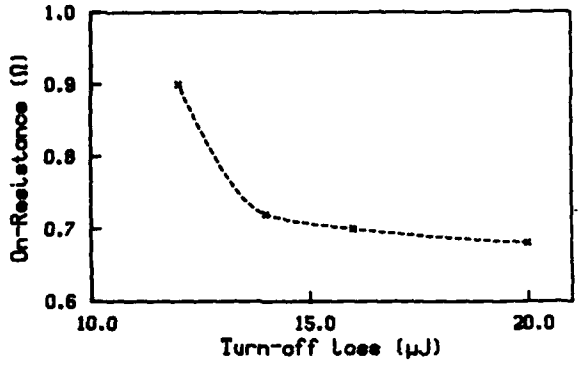


FIGURE 7

On Resistance/Turn-off Loss Trade-off Curve for a Resistive Load

5. CONCLUSION

A compromise between forward voltage and turn-off losses must always be made with a device in the IGBT family. Previously reported methods of controlling the electron-hole plasma have consisted of lifetime reduction and n^+ buffer layers.

In the shorted anode device we have another very simple control method, namely the variation in the degree of n^+ anode short. The level of injected plasma may be decreased simply by designing in a greater percentage of n^+ short. Fig.6 shows the change in device on resistance with different p^+ amounts. This variation in p^+ to n^+ area leads to a trade-off curve between on-resistance and turn-off loss, Fig.7.

REFERENCES

[1] Russell, J.P., Goodman, A.M., Goodman, L.A. and Neilson, J.M., IEEE Electron Device Lett. EDL4 (1983) pp.64-65

[2] Baliga, B.J., Adler, M.S., Love, R.P., Gray, P.V. and Zommer, N.D., IEEE Trans. Electron Devices ED31 (1984) pp.821-828

[3] Nakagawa, A. and Ohashi., IEEE Electron Device Lett. ED6 (1985) pp.378-380

[4] Chow, T.P. and Baliga, B.J., IEEE Electron Device Lett. EDL6 (1985)

Design of a 1600 V Power Bipolar Mode FET (BMFET)

P.Spirito, G.Vitale
Dept.Ingegneria Elettronica University of Napoli, via Claudio 21,
80125 Napoli - Italy

G.Busatto
IRECE - CNR, via Claudio 21, 80125 Napoli - Italy

G.Ferla, S.Musumeci
S.G.S. Microelettronica S.p.A., stradale Primosole 50,
95121 Catania - Italy

The design, the fabrication and the characterization of a BMFET (Bipolar Mode FET) with a maximum voltage of 1600 V and maximum current of 5 A (at h_{FS} = 2.5) are reported. With the help of theoretical models, the effects of physical and geometrical parameters on the performance of the device are investigated and the rules to be used in the design are defined. The switching characteristics of the fabricated devices show that the BMFET gives full Reverse Safe Operating Area and fall-time down to 30 ns on inductive load.

1. INTRODUCTION

The modern families of power devices, used in switching applications, are designed with the goal of reducing power losses and improving reliability. As a consequence new structures, which overcome the drawbacks related to power MOSFET's and Bipolar Junction Transistor (BJT), have been developed. In fact, while power MOSFET's show better performances in term of switching times, Safe Operating Areas and ease of drive as compared with BJT's, they introduce a larger on resistance than the BJT. In recent years devices like COMFET or IGT [1] have been introduced that combine MOSFET operation with BJT operation.

The BMFET too can be considered as a composed device in which a power JFET is combined with bipolar operation, the latter being obtained by forward biasing the gate junction. With proper design, the BMFET can be fabricated as a "normally-off" device in which no gate reverse bias is needed to ensure the channel pinch-off up to the maximum drain voltages [2,3].

In this paper it is reported the design and the realization of a BMFET with a 1600 V sustaining voltage, designed to be used in TV applications. Figure 1 exhibits the schematic

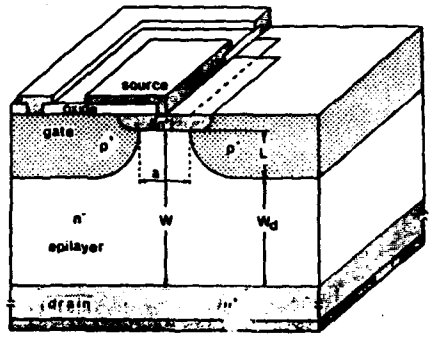


FIGURE 1
Schematic of the elementary cell of the BMFET.

of the elementary cell showing that it consists of a thin N^+ stripe region (source) surrounded by a deep P^+ ring region (gate). In order to obtain a normally-off operation at high voltages with a reasonable geometry of the channel the device has been realized on a very low-doped (>200 Ohm·cm) epilayer 120 μm thick. The required current is obtained by paralleling about 1000 elementary cells over a chip area of 200*200 $mils^2$. A triple guard ring structure ensures the proper high voltage termination at the periphery of the device.

A picture of the complete device before bonding and packaging is displayed in figure 2

The work was supported by Italian National Research Council under the program: "Materiali e Dispositivi per l'Elettronica a Stato Solido".

that shows how the large number of small cells requires the use of a planar LSI technology.

In the following sections the design of the device will be described showing how the geometry of the elementary cell and the doping



FIGURE 2
Picture of a 1600 V 5 A BMFET on wafer.

of the epilayer can be determined in order to satisfy the voltage requirements. The role of the carrier transport parameters in defining the on characteristics will also be investigated. Then the switching behavior will be presented and compared with a BJT of similar ratings.

2. STATIC BEHAVIOR

2.1 The off-state

In the off state, the potential barrier, present along the channel between the gate diffusion, prevents the flow of the current in the source-gate circuit. If the doping of the epilayer is sufficiently low and the geometry of the elementary cell is well dimensioned, the simple built in potential of a PN gate junction is enough to obtain a barrier higher then -0.3 V, needed to originate a drain current less than 10 μa [4].

As shown in figure 1, because of the cylindrical shape of the gate diffusion, the width of the channel results variable along the vertical direction thus generating a strongly two-dimensional potential distribution. The design of the channel geometry has been carried out by using a two-dimensional analytical model [3,4], in which the gates are considered squared and a corrective factor has been used to determine the effective length of the channel. The use of this model simplifies the design of the device and is sufficiently accurate for design purposes without the need of a large calculation times required by a numerical accurate two-dimensional program. The results are reported in figure 3 where, for fixed epilayer doping ($N_{epi} = 1*10^{13}$) and channel width (a = 2 μm), the sustaining voltage is plotted as a function of the channel length for various values of the epilayer thickness. It is interesting to note that the blocking voltage increases with the increase of the channel length and it is limited by the

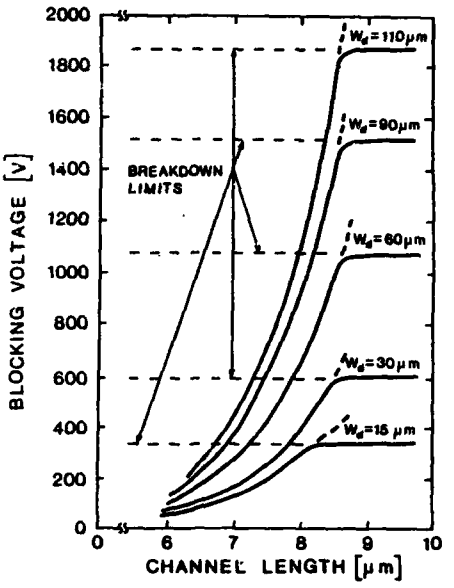


FIGURE 3
Blocking Voltage of BMFET as a function of channel length for various epilayer thicknesses.

breakdown of the gate-drain junction. Hence, once the epilayer thickness has been selected for the maximum voltage to be sustained, the length L of the channel is chosen in the flat zone of the curve as close as possible to its edge in order to get free from the larger sensitivity introduced by the channel length. By using a complete two-dimensional simulation, a final check of the design was made. Its results are not reported for brevity but they show that the potential distribution has a peak higher than -.3 V up to a drain voltage of 1600 V.

2.2 On state

The current amplification of the BMFET is a consequence of two different phenomena: the modulation of the potential barrier into the channel and the conductivity modulation effect into the epilayer, both due to the forward biasing of the gate. Modulation of the potential barrier takes place in the range of

low gate bias and low drain currents and results in a current gain which increases with the gate current. Conductivity modulation effect takes place in the range of high currents, thus determining the maximum operating current of the device. This effect is related to the injection of minority carriers from the gate region and to the correspondent accumulation at the NN^+ source transition. As a consequence, a plasma region is originated in the epilayer below gate and source. The following equation [5], expresses the current gain h_{FS} in the range of high current densities:

$$h_{FS} = \frac{4\, q\, D_n\, A_D}{(A_S \cdot S_S/N_D + A_G \cdot S_G/N_D)\, I_D\, w^2} \qquad (1)$$

where: D_n is the diffusion constant of the electrons in the epilayer, w is the epilayer thickness, I_D is the drain current, A_D is the drain area, A_S (A_G) are the source (gate) areas, S_S (S_G) are the recombination velocities of the source (gate) layer, defined, according to [6], as:

$$S_S = \frac{D_{ps}N_D}{L_{ps}N_{DS}} \mathrm{Exp}\left(\frac{\Delta V_S}{V_T}\right) \mathrm{Coth}\left(\frac{W_S}{L_{ps}}\right) \qquad (2a)$$

$$S_G = \frac{D_{ng}N_D}{L_{ng}N_{AG}} \mathrm{Exp}\left(\frac{\Delta V_G}{V_T}\right) \mathrm{Coth}\left(\frac{W_G}{L_{ng}}\right) \qquad (2b)$$

where: W_S (W_G) is the source (gate) thickness, D_{ps} (D_{ng}) are the source (gate) diffusion coefficient, L_{ps} (L_{ng}) are the source (gate) diffusion length, N_{DS} (N_{AG}) is the source (gate) doping, ΔV_S (ΔV_G) account for the effects of band gap narrowing due to the high doping of the source (gate) layers.

These expressions for S_S and S_G have been obtained for the case of high injection, then the contribution due to the space charge region [7] has been neglected.

Observing the equation (1), it is important to note that h_{FS} is inversely proportional to S_S and S_G. Moreover, the effect of the source recombination is much smaller than that of the gate recombination because the source area is smaller than the gate area in order to satisfy the design requests for the off state. As a consequence, the doping parameters of the gate region can be determined in order to optimize the current gain of the device. By using equation (2b), a plot of the gate recombination velocity, S_G, as a function of the gate doping for an epilayer doping of 10^{13}, has been determined and reported in figure 4. The plot shows a minimum for S_G

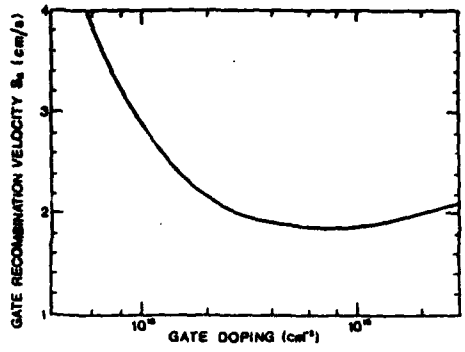


FIGURE 4

Electron recombination velocity in the gate as a function of the gate doping.

at a gate doping of $6 \cdot 10^{18}$. Moreover the curve is quite flat near the minimum thus a very small sensitivity has to be expected due to the gate doping on the h_{FS}.

3. Switching behavior

The BMFET has the possibility to assist the turn off with the extraction of minority carrier from the gate junction, by reverse biasing the gate terminal. As a consequence the BMFET is much faster than the MOS-gated bipolar devices in which the reduction of the minority carrier stored in the epilayer is left only to the internal recombination.

As compared to the BJT, the BMFET is faster as shown in figure 5 where the a plot of the

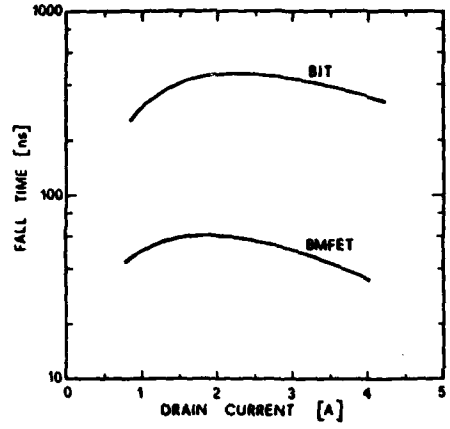


FIGURE 5

Fall time for switching on inductive load of BMFET compared with BJT, as a function of the output current.

fall time is reported as a function of the drain current . The figure refers to the fall time observed during switching on inductive load for both a BJT and a BMFET with same characteristics. In the whole range of drain current the BMFET exhibits a fall time one order of magnitude lower than that of the BJT with a minimum value of 30 ns observed at the maximum drain current.

The reasons are related to the fact that, when turned off, the BMFET ends in a situation in which the current flow is controlled by a potential barrier in the channel. That means that after an initial bipolar transient, during which the excess of minority carriers is swept out through the gate, a second phase takes place which is essentially unipolar and controlled by gate-emitter and the gate-drain capacitances. During this latter part of the transient, the switching of the BMFET is similar to that of a MOS device and hence very fast. The incidence of this latter portion is dependent on the load, in particular for switching on inductive load, in which the drain current is initially constant and the voltage varies slowly, the bipolar transient ends before the fall-off of the current that consequently goes down during the unipolar transient thus being very fast. Moreover also the bipolar transient of the BMFET is faster than the BJT because of the absence of any distributed base resistance that limits the extraction of minority carriers in BJT.

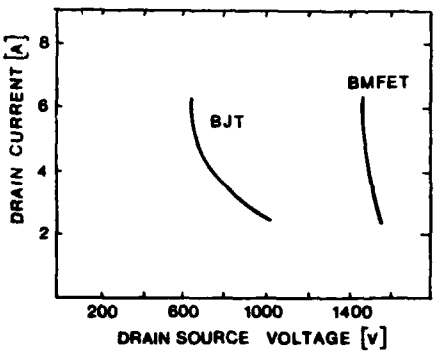


FIGURE 6
Reverse Safe Operating Area of the BMFET compared with BJT.

The comparison between BJT and BMFET is favorable to the latter also with regard to the reverse safe operating area, as shown in figure 6. In fact, while the sustained voltage of the BJT reduces, at high current, to about one half of the maximum voltage, the BMFET sustains the maximum voltage up to the maximum current. The reason of such a behavior is that, for the BMFET, the gates are disposed laterally to the conduction channel, so there is not the feedback effect which limits the maximum voltage of the BJT. On the other hand no emitter crowding effects take place and consequently second breakdown, due to the nonuniformity of the current, is not observed.

4. Conclusion

The paper has explained how the design of a BMFET with prescribed on and off characteristics can be performed on the basis of simple but accurate analytical models. Moreover it has been demonstrated that a BMFET with sustaining Voltage up to 1600 V can be realized and the full voltage can be sustained without reverse biasing of the gate. The devices fabricated fit very well with the theoretical predictions with regard to both the off and on performances.

Moreover the switching characteristics of the fabricated devices are reported, showing that because of the structure of the device the BMFET is substantially faster than the BJT and the fall time is comparable with that of power MOSFET.

Finally, the paper shows that the BMFET exhibits a full reverse Safe Operating Area allowing the sustaining of the maximum voltage up to the maximum current.

REFERENCES

[1] Russel, J.P., Goodman, A.M., Neilson, J.M., IEEE Electron Device Letters (Vol. EDL-4, 1983) pp. 63-65.

[2] Caruso, A., Spirito, P., Vitale, G., Busatto, G., Ferla, G. IEEE Power Electronics Specialist Conference (Canada, Vancouver, 1986) pp.

[3] Caruso, A., Spirito, P., Vitale, G., Busatto, G., Cocorullo, G., Ferla, G. IEEE Industry Application Society Meeting (Colorado, Denver, 1986) pp.

[4] Ohmi, T. IEEE Trans. Electron Devices (Vol. ED-27, 1980) pp. 536-545.

[5] Bellone, S., Caruso, A., Spirito, P., Vitale, G. Solid-State Electronics (Vol. 26, 1983) pp. 403-413.

[6] Shina, A., Chattopadyaya, S.K. IEEE Trans. Electron Devices (Vol. ED-25, 1978) pp.1412-1414

[7] Ram, G.V., Tyagi, M.S. Solid-State Electronics (Vol. 24, 1981) pp. 753-761.

MODELLING BIPOLAR TRANSISTOR SECOND BREAKDOWN DURING TURN-OFF BY SOLUTION OF THE FUNDAMENTAL DEVICE EQUATIONS

S.A. HIGGINS, M.K. JOHNSON, P.A. GOUGH, J.A.G. SLATTER

Philips Research Laboratories, Redhill, Surrey, U.K.

Numerical simulations have been made of the turn-off of two-dimensional power bipolar transistor structures, under inductive loading, by alternate solution of the fundamental device equations and the circuit equations at each time step. The results have shown that current spreading reduces the magnitude of the electric field in the collector produced by the space charge of mobile electrons, so leading to improved reverse bias second breakdown performance. Removal of the central portion of the emitter effectively increases the current spreading and gives a further reduction in the space charge induced field.

During the turn-off of a power bipolar transistor, under inductive loading, high current densities occur in the transistor as conduction is squeezed towards the centre of the emitter. The large current densities can cause a space charge of electrons to appear in the lightly doped collector (assuming an n-p-n transistor) where a high electric field and avalanche generation of carriers can result, leading to second breakdown [1-2]. This effect limits the Reverse Bias Safe Operating Area (RBSOAR) of the transistors and has proved difficult to model accurately in the past by effectively one-dimensional methods. To model current squeezing, the two-dimensional, time dependent distribution of carriers and potential in the device must, as a minimum, be considered.

We have modelled the turn-off of a typical power switching bipolar transistor under inductive loading, taking into account two-dimensional effects. A cross section of the device normal to the emitter finger diffusion pattern is shown in figure (1) and the circuit is shown in figure (2).

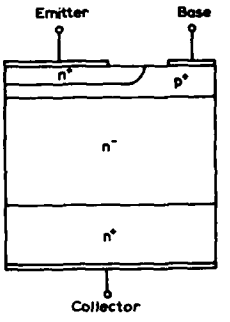
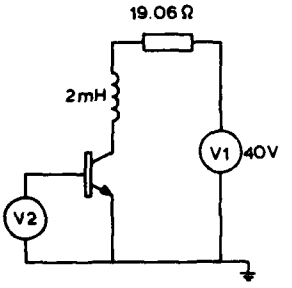


FIG. 1. Cross section of modelled transistor



V2 = linear ramp from 1V to 0V
for 0 < t < 200ns
= 0V for t > 200ns

FIG. 2. Inductive switching circuit

We have modelled this two-dimensional transistor (1 cm in length) assuming, by symmetry and uniform conduction over the fingers, that this represents the whole transistor. The internal distributions of carriers and potential are found by solving the two-dimensional, time dependent, transport and continuity equations for electrons and holes, and Poisson's equation, on a rectangular mesh, by Newton's method. Alternate solutions of these equations and the collector circuit equation leads to collector current and voltage values which at each time step satisfy the collector circuit equation. Time integration accuracy is checked by comparing the results from a time step with that from two time steps of half the size; if the accuracy criterion is not met the solution is recalculated with a reduced time step [3].

In the simulations the transistor was turned-off by reducing the base-emitter voltage from its on-state value of 1 volt to 0 volts, in a linear ramp, over a time period of 200ns. The computed base current, collector current, and collector voltage are shown in figure (3), where it can be seen that it has been possible to model the device at collector voltages greater than 1kV, with dV/dt values of ~10kV/μs, during turn-off.

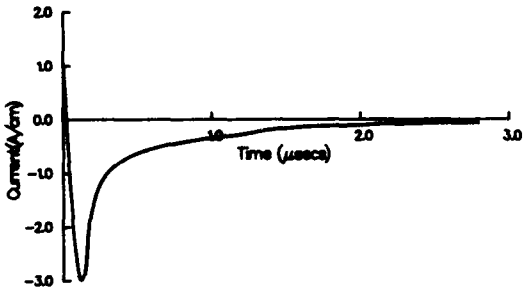


FIG. 3(a). Base Current during Switch-off

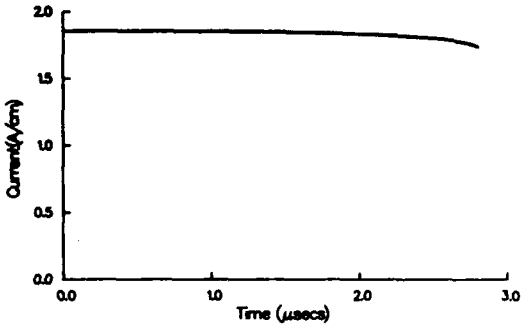


FIG. 3(b). Collector current during switch-off

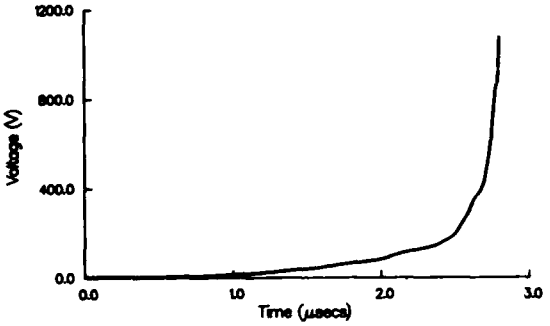


FIG. 3(c) Collector voltage during switch-off

The variation of the current distribution, during the switch-off process is illustrated in figure (4). The last graph in the sequence shows the final extent of the current crowding towards the centre of the emitter and the spreading out of the current as it passes across the collector. The final distribution of the electron density and electric field, at t=2.8 μs, Vce=1070 Volts, is given in figure (5)

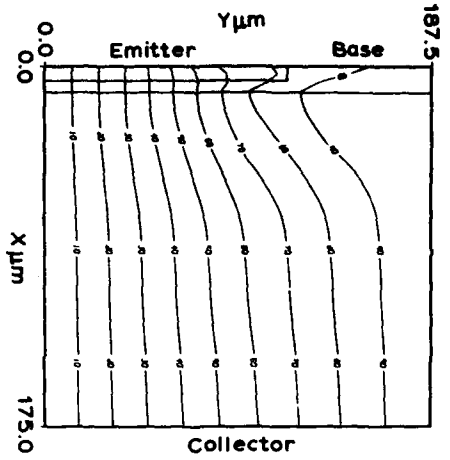


FIG. 4(a) Current distribution at t=1.0 μs

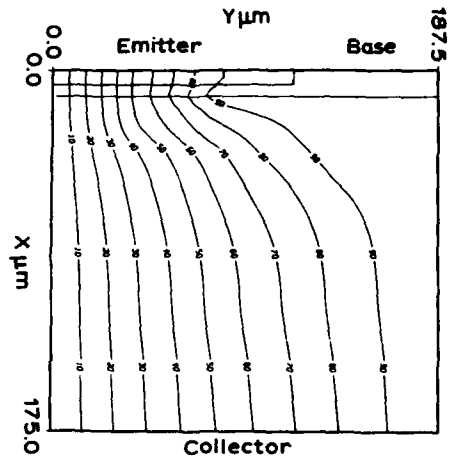


FIG. 4(b) Current distribution at t=2.0 μs

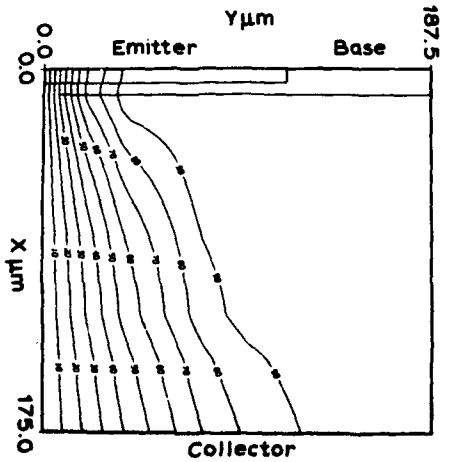


FIG. 4(c) Current distribution at t=2.8 μs

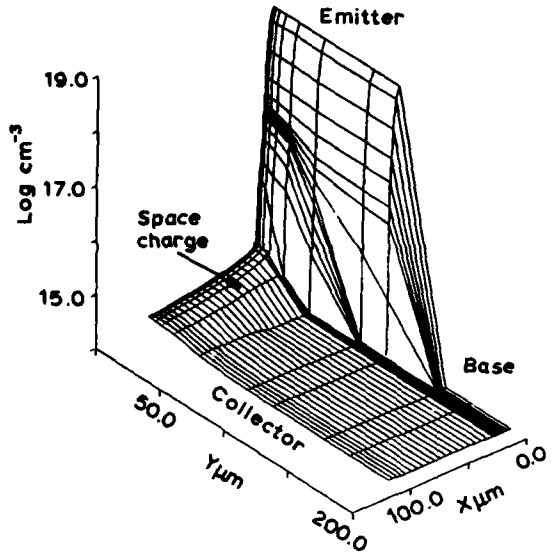


FIG. 5(a) Electron density at t=2.8 μs

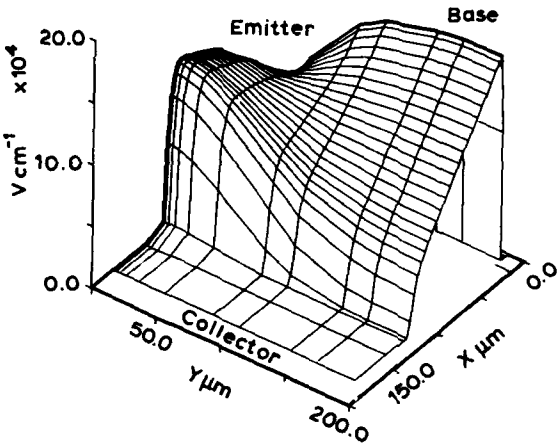


FIG. 5(b) Electric field at t=2.8 μs

These results confirm previously reported simulations [4] in showing that current spreading in the collector significantly reduces the value of the peak electric field under the centre of the emitter caused by the electron space charge. In the device modelled, with a comparatively thick collector, the current spreading also moves the position of the peak field from the n^-/n^+ junction into the interior of the collector. The results also show that for the given bias conditions the

field under the centre of the emitter, caused by the space charge of electrons, does not reach a value which will cause a significant multiplication of carriers, even at Vce=1kV.

We have also modelled a two-dimensional structure which differs from the previous structure in that the central half of the emitter diffusion has been masked out, and which we shall refer to as a hollow emitter structure. Starting from the same on-state conditions, we have turned this device off with the same base drive. Figure (6) shows the current spreading and figure (7) illustrates the electric field distribution at t=2.27 μs, Vce=650V. We find that the effective offset of the emitter produces a greater current spreading compared with the structure discussed previously, the effect of which is to reduce the space charge and the peak field in the lightly doped collector for a given collector voltage.

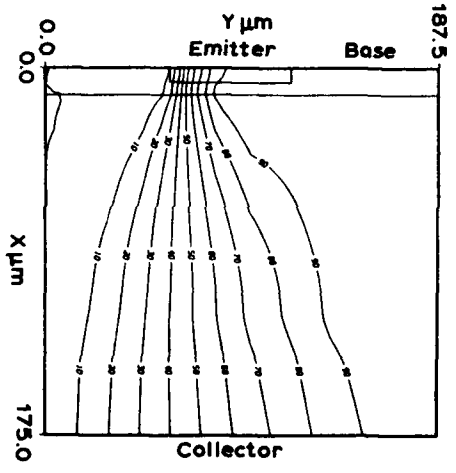


FIG. 6. Current spreading in hollow emitter structure at t=2.27 μs, I_c=1.8A, Vce=650V

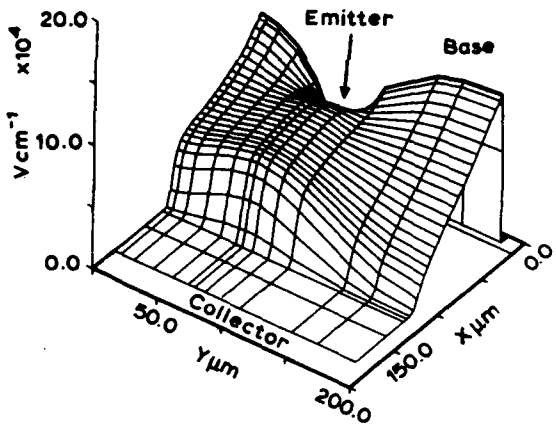


FIG. 7. Electric field in hollow emitter structures at t=2.27 μs, I_c=1.8A, Vce=650V

In figure (8), we show the peak electric field in the space charge region at a given collector voltage for both the solid emitter structure discussed previously and the hollow emitter structure. These results suggest that bipolar transistors with hollow emitters will have superior reverse bias second breakdown performance.

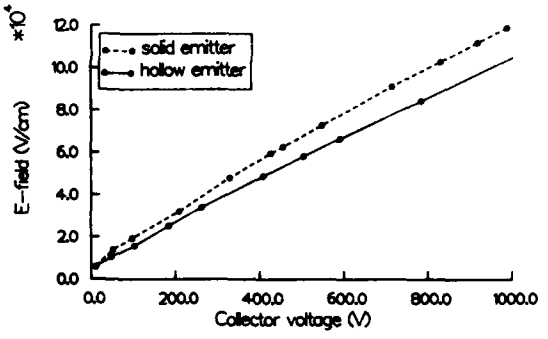


FIG. 8. Comparison of solid and hollow emitter structures

In conclusion, the simulations in two-dimensions of these bipolar transistor structures show that they have good second breakdown properties during turn-off and that there are advantages to be gained from changing the emitter structure.

REFERENCES

[1] Schrafft H.A., "Second Breakdown - A Comprehensive Review", Proc. IEEE, Vol.55, pp.1272-1288, 1967.

[2] Hower P.L., Reddi V.G.K., "Avalanche Injection and Second Breakdown in Transistors", IEEE Trans. EI. Dev., ED-17, pp.320-335, 1970.

[3] Gough P.A., Johnson M.K., Higgins S.A., Slatter J.A.G., Whight K.W., "Two-Dimensional Simulation of Power Devices with Circuit Boundary Conditions", Nasecode 5, Dublin, 1987.

[4] Hwang K., Navon D.H., Tang T., Hower P.L., "Second Breakdown by Two-Dimensional Numerical Analysis of BJT Turn-off", IEEE Trans. EI. Dev., ED-33, pp.1067-1071, 1986.

PROCESSING AND CHARACTERIZATION OF ULTRA-SMALL SILICON DEVICES

G. A. Sai-Halasz

IBM, T. J. Watson Research Center, Yorktown Heights, N.Y. 10598

Processing, design, and characterization issues are discussed for advanced field-effect (FET) and bipolar transistors. Results are presented from work on N-channel FET's with gate lengths below $0.1\mu m$, and on the role that polysilicon emitter contacts play in high current, high speed bipolar devices. For FET's over 750mS/mm transconductance was achieved at liquid nitrogen temperature operation. In the case of bipolar devices it was found that maximizing the gain enhancement derived from the polycrystalline/single-crystal interface without regard to resistive effects does not lead to the highest performance in submicron transistors. Based on previous experience and on our recent work, we believe that silicon technology faces no insurmountable obstacles as it progresses into the deeply submicron regime.

1. INTRODUCTION

Silicon bipolar and field effect transistors, FET's, can look back to an extremely successful three decades. By now, if we judge by the expended research and development efforts, silicon technology must be one of the most mature ones in existence. Because of this it is often assumed that it ran its course, and efforts in the direction of performance or density improvements give ever diminishing returns. In reality silicon technology marches forward in a relentless fashion. Minimum dimensions, typically characterized by gate length in FET's and base width in bipolars, are continually shrinking. Although the rapid device performance improvements of the 1960's slowed significantly by the 1970's, integration levels and system performance continue to improve. Still, questions are being raised regarding the limits of continually scaling down dimensions. In this paper we concentrate on specific design, processing and characterization issues encountered as a result of shrinking dimensions. For FET's we do this in the light of a recent effort in our laboratory to investigate the feasibility of a self-aligned FET technology in the $0.1\mu m$ gate length regime. In the case of bipolars we concentrate on the specific issue of resistance due to polysilicon emitter contacts, an unanticipated problem that arose as a consequence of decreasing dimensions. Finally, we'll briefly comment on a few selected topics.

2. SMALL FET's

2.1. Work in the submicron environment

Those who enter the world of deeply submicron FET devices have mainly bad news to contend with. Classical FET scaling [1] runs into a variety of nonscaling parameters and associated detrimental effects. Among others, these include mobility degradation, [2] inversion-layer broadening, [3] tunneling through the gate insulator, and in general the onset of velocity saturation. But there may be advantageous effects awaiting as well. Such are the possibility of velocity overshoot in very short devices, [4] and the weakening of those detrimental effects that are associated with an energy threshold. In this latter class belong avalanche breakdown and hot carrier injection into insulators. Apart from novel effects, the worth of shrinking dimensions hinges on such practical difficulties as linewidth and alignment control, thin insulator reliability, shallow junction fabrication, the limitations of contact and spreading resistivities, and many others. Although

theoretical treatments abound on the perceived limits of Si FET technology, relatively little experimental work has been performed below the 0.5μm regime. It is this area that our work [5] was concentrated upon.

The general aims of our work were, (1) to study scaling parameters down to the 0.1μm level, (2) to fabricate a test vehicle in as conventional a manner as possible, that then can serve as a starting and reference point for radical departures in processes and/or device concepts, (3) to see what transconductance and switching times are attainable in such a technology, and (4) to investigate what, if any, novel effects are observed at these dimensions.

2.2. Design and test structure description

The junction potential of silicon sets a lower limit on operating voltages independently of device dimensions. Consequently, one might be forced to work at higher voltage levels than dictated by ideal scaling. This is most detrimental once electric fields reach levels where velocity saturation dominates device behavior. In such an environment, raising operating voltage level leads only to increased power consumption without an accompaning increase in performance. Thus one must strive for operating at the lowest possible voltage level. These considerations naturally lead to liquid nitrogen (LN_2) temperature FET operation. At LN_2 temperature the main obstacles to lowering the threshold, namely subthreshold conduction and threshold shift due to temperature change, are drastically reduced. There are other advantages to LN_2 temperature FET operation as well: improved performance, mainly arising from lower line resistances and better punch through behavior. Also at 77°K one can forward bias the substrate. This bias counteracts the junction potential which otherwise prevents the scaling down of the depletion regions around the junctions. [6] In our view, although there are no fundamental obstacles to a room temperature 0.1μm gate length FET design and operation, considering the difficulties in processing, to make the performance worthwhile of the fabrication effort one has to go to reduced temperature operation.

Although ultimately one is aiming at CMOS, as a first feasibility study at the 0.1μm gate length level our efforts were directed toward NMOS. The lowest threshold regarded to be practical even at LN_2 temperature was 150mV, the same as proposed for a 0.25μm LN_2 temperature design.[6] This design path leads to a 0.6V power supply. The optimum choice for substrate bias would be as large as possible while still avoiding significant leakage currents. A 0.6V forward bias on the substrate, the same as the drain voltage, is a reasonable compromise, and saves an additional power supply. Because of freezeout at LN_2 temperature makes the design of depletion mode devices difficult, and because of the attractiveness of eventually developing CMOS, no attempt has been made to fabricate depletion-mode devices. Instead, enhancement-mode loads with a separate gate voltage supply were utilized. For circuits operating in the velocity saturation regime the optimal value for the ratio of the widths of active and load devices was found to be 4:1.

Seven different chips were assembled with a variety of test structures. Three of the chips contained inverter chains. The other four chips served to house various parametric test sites, like capacitors, very large devices, linewidth, alignment and bias monitors, contact and sheet-resistivity measurement sites. Simple circuits were also built to measure the actual capacitance of the small gates during device operation, to observe gate leakage, and to measure intrinsic transconductance and device noise. Many of these test structures were used earlier in our laboratory in their larger versions. [7] On most chips the test sites have been repeated several times to give versions with different gate lengths at the design level. The actual fabricated gate lengths on the chips ranged from a maximum 0.25μm down to a minimum of 0.07μm. The number of sites were limited by the output pads that could be conveniently located around the high resolution field. To convey some of the complexity of these nominally 0.1μm test chips, Fig. 1 shows the picture of one of them. In such pictures almost the only visible thing is the metallization. To observe the finer features we had to remove the covering layers to expose the gate level. Figure 2 shows an SEM micrograph of 0.07μm long gates in an inverter following such stripping.

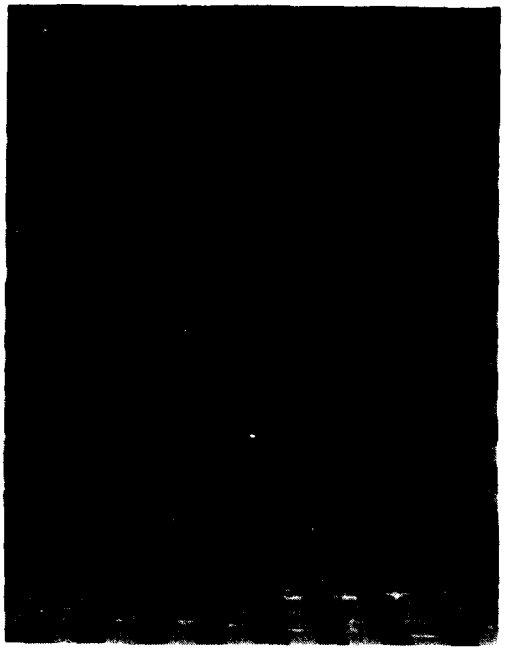


FIGURE 1
Picture of a chip containing inverter chains and support circuits. The high resolution area is 0.25x0.25mm².

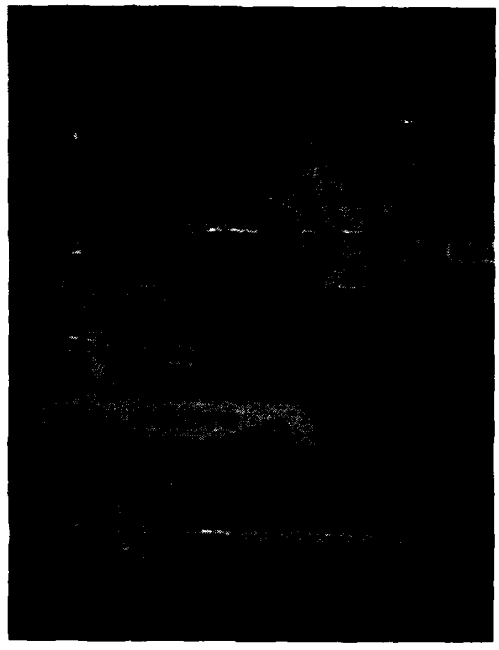


FIGURE 2
SEM micrograph of an inverter after stripping the metal and conformal oxide. Diffusion and field regions, marks due to contact holes, and nitride sidewalls are visible together with 0.07μm long gates.

The schematic cross section of a nominal device is shown in Fig. 3.

2.3. Processing

Our NMOS process called for five lithographic levels, including the one for alignment marks. They were all done by direct write electron-beam. The electron-beam system had the capability to produce a few tens of nm features on the wafer, with an overall overlay accuracy of 50nm in a 0.25x0.25mm² field. [8] In some cases the overlay turned out to be as good as 10nm.

Processing at these dimensions leads to unavoidable difficulties. Consequently, whenever it was possible, established process steps were used with added modification if needed. In this spirit a conventional semi-recessed oxide isolation was used. However, this type of isolation allowed only 0.25μm ground rules for the diffusion level. Vertical scaling and doping levels were attempted to be fully consistent with a 0.1μm gate length. This was only relaxed for the gate oxide, where full vertical scaling would have called for a 2 to 2.5nm thickness. Considering that we were making a first attempt and were facing many unknowns, we increased the oxide thickness to 4.5nm on most wafers for reliability reasons. (Tunneling would have allowed the use of thinner insulator because our voltage levels were

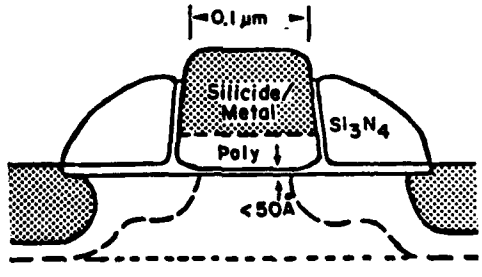


FIGURE 3
Schematic cross section of the NMOS device. Long dashed line indicates junction edges, short dashed line shows the peak of the boron implant.

scaled down.) As mentioned earlier, the actual gate lengths spanned a considerable range on each chip. The doping levels were chosen to be consistent with the design for the 0.1μm gate length devices, i.e. giving the threshold, substrate sensitivity, punchthrough control correctly for this gate length. On each wafer the seven chips were reproduced three times, giving a total of 21 per wafer. Because of the uncertainties as to which process detail will lead to optimal devices, many process variations were used. The main process steps, with the reasons behind them and the significant experimental variations were the following:

1. The wafers, 2Ωcm p-type, received a backside boron implant to provide for good contact even at LN_2 temperature.

2. Patterned, and fabricated by liftoff $TaSi_2$ alignment marks. This material has sufficiently coarse grain also to provide focusing targets. The alignment marks had minimum dimensions of 0.25μm. Four marks were provided in each corner of the 250μm chip field so that if any marks were damaged three spare ones were still available. Marks in the corners of the chip field could be scanned simultaneously, to adjust the offset, field size, rotation, and orthogonality of the exposure field.

3. Grew 10nm pad oxide and deposited 100nm nitride. Patterned the diffusion regions and removed the nitride-oxide stack from field regions. Implanted boron for isolation. Grew 160nm dry, semi-recessed field oxide at 950°C.

4. Stripped nitride-oxide stack from diffusion regions. Implanted boron for threshold control. Doses ranged from 2.5 to $5.5x10^{12}$ /cm². Two implant energies, 15 and 30keV, were used to provide a reasonable trade-off between process sensitivity, short-channel effects, and performance. The aim was that for a given thermal budget the surface B concentration be as low as possible, while peaking the dopant below the surface for punchthrough control.

5. Grew thermal SiO_2 gate insulator ranging in thickness from 3.3nm to 4.5nm, at 800°C in dry oxygen with HCl.

6. Deposited 100nm thick undoped polysilicon for gate.

7. Patterned gate level, and lifted off a metal etch mask. The most critical step in terms of resolution was this one. It was patterned using a double layer PMMA resist. After metal *lift off, the polysilicon was* reactive ion etched.

8. Grew 7.5nm oxide on polysilicon at 850°C in dry oxygen with HCl.

9. Implanted source/drain extensions, [7] which serve the purpose of reducing short channel effects. In different splits arsenic and antimony [9] were utilized as dopant species. Doses were in the few times 10^{14} /cm² range, with energy down to 10keV. Multiple-energy implants were used in some cases to achieve a more "boxlike" profile. The profile at the junction edge is related to device performance. [7,9] The source and drain junction must be as abrupt as possible, since the main resistance component leading to transconductance degradation arises in the region where the source/drain dopant concentration gradually drops toward the channel. The choice of antimony for source/drain extension is dictated by the fact it is more suitable for achieving abrupt junctions. [9]

10. Deposited 100nm nitride for sidewall spacer, and reactive ion etched the nitride.

11. Implanted arsenic source/drain. Doses were 2 to $5x10^{15}$ /cm², at 20keV.

12. Activated source/drain. Furnace thermal budgets ranged from 900°C, 30 minutes to 850°C, 20 minutes. Some wafers received rapid thermal anneals of 1050 and 1100°C for 10". The final depth was estimated to be ~50nm for the source/drain extensions and 100nm for the "deep" junctions.

13. Self-aligned metallization formed on gate and source/drain for most wafers.

14. Deposited 100nm low-temperature conformable oxide overlay.

15. Patterned contact hole level, and reactive ion etched contact opening.

16. Patterned metal level via lift-off of 300nm Ti/Al, and annealed in partial hydrogen ambient.

2.4. Characterization

Of the tested wafers ~80% of the structures were operational. This percentage includes some sites that depended on the operation of many devices,

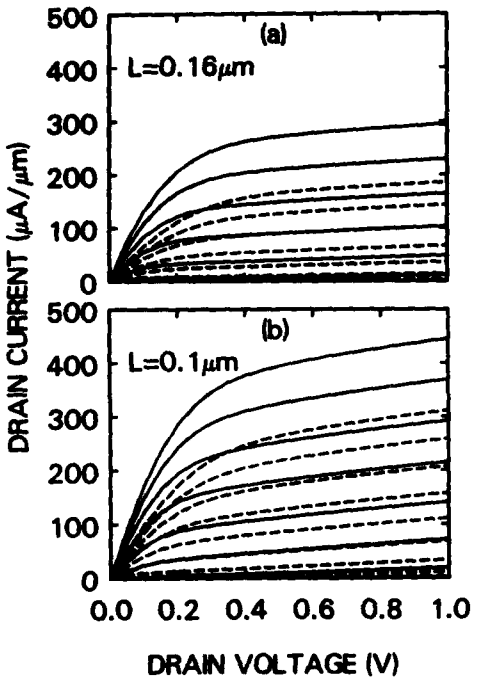


FIGURE 4

Characteristics of 0.16μm gate length (a), and 0.1μm gate length (b) devices from the same wafer. Solid lines are for LN_2 and dashed lines for room-temperature operation. Gate voltage increments are in 0.1V steps up to 0.8V. Substrate bias is 0V at room and 0.6V at LN_2 temperature. The LN_2 thresholds are ~0.1V above the room temperature ones. Gate oxide is 4.5nm thick. The wafer received a channel implant recessed from the interface, and has antimony source/drain extensions. At 77°K the 0.1μm gate length device has a g_m of 760mS/mm and a G_m of 440mS/mm.

such as inverter chains with on-chip output amplifiers. Considering that at present many of the utilized process steps were of an experimental nature, this is quite satisfactory yield.

The first point of importance is basic device behavior. Figure 4 shows the drain characteristics of two devices from the same wafer, with different gate lengths, at both room and LN_2 temperature. One can see that the device characteristics in the scaled down voltage regime appears almost "long-channel like", in that the output conductance is low. However, the equal spacings of the saturated currents with equal gate voltage steps is the main feature of the devices, indicating strong velocity saturation. Notwithstanding this fact, the devices have exceptional transconductances. In the saturation regime at 77°K the 0.1μm gate length device has a small signal transconductance, g_m , of 760mS/mm. For estimating performance in dense circuits the large signal transconductance, G_m , (current at the design values of $V_d = V_g = 0.6V$ divided by the designed 0.6V maximum gate drive) is the more relevant parameter. [10] This is 440mS/mm for the same device. To our knowledge [11,12] both of these transconductance values are the highest measured to date for Si FET's. They compare quite favorably with competing high-performance technologies.

If we take current output for the same gate drive as measure of device quality, the LN_2 temperature operation resulted in approximately 50% improvement for the 0.16μm gate-length device and ~40% for the 0.1μm one. (Note that the room temperature thresholds are ~100mV below the LN_2 ones.) Two things are somewhat surprising about the LN_2 versus room temperature device operation. First, the improvement due to low temperature is more than expected. This would point in the direction of some overshoot phenomena. On the other hand, and this the second puzzling aspect because it contradicts the conclusion of the first, the shorter devices improve proportionately less then the longer ones upon cooling, arguing against velocity overshoot.

Overall device features were nearly, but not exactly, like those predicted by earlier drift/diffusion finite-element, FIELDAY [13], modeling on which the design was based. This is illustrated in Fig. 5, where the comparison is shown of the 0.1μm gate length device of Fig. 4 with the FIELDAY simulation that was done at an earlier time, when the project was still in the device design stage. In the simulation the mobility is too high, giving too much current at low drain voltages. It is not difficult to match exactly the experimental curves with FIELDAY for any of the gate lengths, with a reasonable choice of parameters. However the same set of parameters does not match all the devices with different gate lengths. The situation becomes much more complicated if we look at wafer to wafer variation in device characteristics.

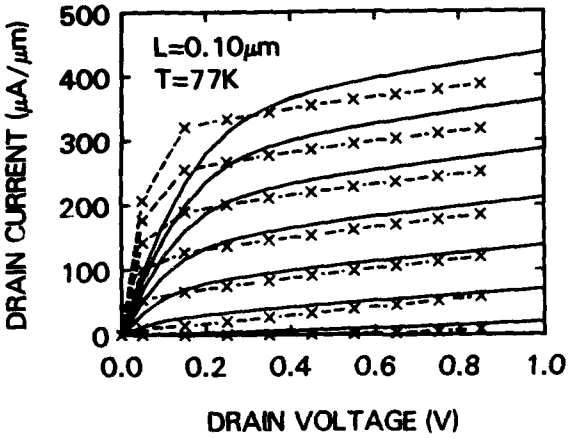


FIGURE 5
Comparison of experiment (solid lines) and simulation (dashed lines and symbols) for 0.1μm a gate length device at low temperature. Gate and substrate biases as on Fig. 4. The simulation is from a drift-diffusion finite element model. (FIELDAY) It was performed prior to device fabrication and served as the basis for the device design. The mobility assumed in the model was too high.

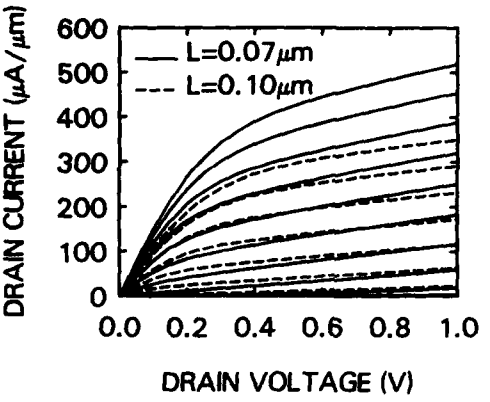


FIGURE 6
LN_2 characteristics of 0.07μm and 0.1μm gate length devices from a wafer processed with different splits from the devices of Fig. 4. Gate voltage increments are in 0.1V steps up to 1.0V. Substrate bias is 0.6V. Gate oxide is 4.5nm thick. These devices show higher output conductance and more resistive effects than the ones in Fig. 3. The figure illustrates the importance of the processing details on device performance. The 0.07μm gate length device has a g_m of 640mS/mm, higher than the 0.1μm device from the same wafer, but lower than the 0.1μm device on Fig. 3., in spite of the longer gate of the latter.

They are extremely sensitive to the details of processing. No such thing was even remotely found as a "generic" 0.1μm gate length device behavior. This is evident by comparing device characteristics shown in Fig. 6 with those of Fig. 4. Figure 6 shows LN_2 temperature device behavior from another wafer having gate lengths down to 0.07μm, made with somewhat different process conditions. While these too are excellent devices, the 0.07μm one having a g_m of 640mS/mm, they show more resistance and an overall lower current output for a given gate drive. When the intrinsic transconductances are estimated for the devices shown in the two figures, the 0.1μm gate length one of Fig. 4 still comes out ahead of the 0.07μm one of the other wafer, in spite of its longer channel length. It was found that high transconductances were associated with low channel doping at the oxide interface, and with antimony source/drain extensions. So far because the strong process dependence of device performance hinders precise modeling, it is not clear to us whether velocity overshoot [12] manifests itself either at room or LN_2 temperature. A more thorough investigation is presently in progress.

Analysis of data for small drain voltages combined with data from large, 100μm channel-length devices gave the following typical parameters at LN_2 temperature. Field effect mobility was approximately 1500cm²/Vsec at turn-on, but deteriorated ~30% by the time the gate voltage reached 1V over threshold. Source and drain resistance for the best wafers was below 130Ωμm per side. Here too, the precise situation for the smallest devices is not quite clear. Figure 7 shows a comparison in behavior of a 100μm and a 0.1μm long device from the same wafer. The figure shows the current normalized for one square of the devices (width equals length) for low drain voltages as function of gate drive. In principle, the behavior should be independent of gate length. This was clearly not the case, the small device fared poorly in the comparison. The main reason was source/drain resistance. However, more detailed analysis showed that source/drain resistance is not the complete picture. The short device had lower mobility than the long one as well. In spite of the appearance given on Fig. 7, the source/drain resistance and mobility of the

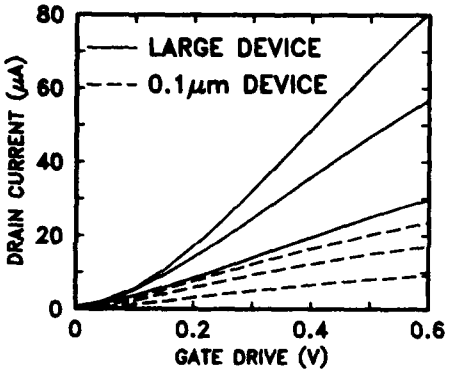


FIGURE 7
Low temperature, low drain bias characteristics normalized to one square (length=width), derived from a 100μm and a 0.1μm gate length device. Drain biases are 50, 100, and 150mV for both devices. The effects of source/drain resistance and somewhat lower mobility is clearly evident in the case of the small device.

small device was well within tolerable limits. According to circuit simulations the parasitic contributions to the overall performance degradation were not too significant when they were compared to the intrinsic degradation inherent in velocity saturation, which dominated at higher drain voltages. The use of LN_2 temperature assured that subthreshold current was ideal down to the 0.1μm level on almost all wafers, in spite of the shallow junctions and the forward substrate bias. It changed a decade for each 27mV of gate voltage change.

2.5. Conclusions

Although a full performance evaluation is not yet complete, we are quite confident that with proper design, and with processing along established avenues, an FET technology at, or even below, the 0.1μm gate length level appears feasible, and performance improvements will accompany the shrinking of dimensions into this regime.

3.BIPOLAR DEVICES

3.1. Scaling issues

Bipolar transistors are evolving at a rapid pace as well. Their scaling [14] is a less straightforward case than that of FET's, but ultimately it is associated with similar issues: thinning the base, decreasing horizontal dimensions, and adjusting doping levels. In processing, possibly the most critical issues are the controllable reduction of base width and the shallowness of the emitter junction. As an example of processing and characterization issue arising with scaling down dimensions we'll discuss the question of polysilicon emitter contacts and their influence on performance. [15]

3.2. Polysilicon emitter contact

Reducing the base-emitter junction in conventional metal-contacted transistors causes the base current to be unacceptably high and very sensitive to the contact properties. Polysilicon emitter contacts significantly improve bipolar device performance, because they produce a highly abrupt, shallow base-emitter profile without a reduction in current gain since the metal contact is moved away from the junction The reduced mobility in the polysilicon [16] and the presence of an interfacial tunneling barrier [17] at the polycrystalline-Si/single-crystal-Si (poly/single) interface result in a reduced base current density by a factor of 3 or more. The poly/single interface, however, not only acts as a barrier to hole transport (base current), but also increases the resistance for electrons (emitter current). [18] Since the largest current density in the device passes through the emitter contact, the specific contact resistivity must be kept as low as possible to prevent a degradation in transconductance. The performance loss due to contact resistivity only shows up in small devices at high current densities. [19] Unfortunately these are exactly the conditions under which high performance logic circuits operate. It is of significant importance to tie in electrical behavior with the nature of the poly/single interface. It is known that the interface structure changes significantly as a result of high temperature annealing and that these changes are enhanced at high doping levels.

3.3. Device fabrication

In order to determine the potential performance trade-off between base current and emitter resistance we compared the structural and electrical aspects of advanced technology bipolar transistors which differed only in the

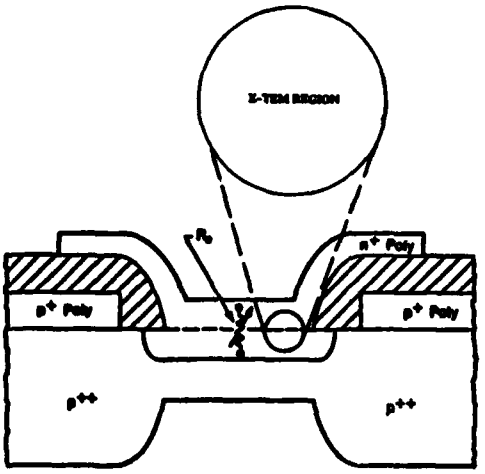


FIGURE 8
Schematic cross-section of double-poly self-aligned bipolar transistor, with the emitter resistance indicated.

annealing treatment of the polysilicon emitter contact. The devices used in our experiment were self-aligned double-polysilicon bipolar transistors. whose schematic cross-section is shown in Fig. 8. The whole process description can be found in [15], here we are only detailing those aspects that had a direct bearing on the poly/single contact interface. The emitter opening was etched using a selective chemical etch, so as to prevent surface contamination that can occur in reactive ion etching procedures. In-situ, As-doped polysilicon was deposited at 650 °C using N_2 as the carrier gas. The emitter surface was given a buffered HF clean immediately prior to this deposition to minimize any residual native oxide on the surface. Two different emitter anneals were used: either 900 °C for 10 min. or 800 °C for 45 min. The second anneal condition was designed to result in approximately the same integrated base doping, while minimizing the base-emitter junction depth. Metallization consisted of a Ti/Al/Si sandwich, which was annealed in a manner known to give extremely low contact resistance between the metal and polysilicon.

3.4. Electrical characterization

Extensive DC and AC characterizations were done of both large, over 1000μm^2, and small, ~10μm^2, emitter area

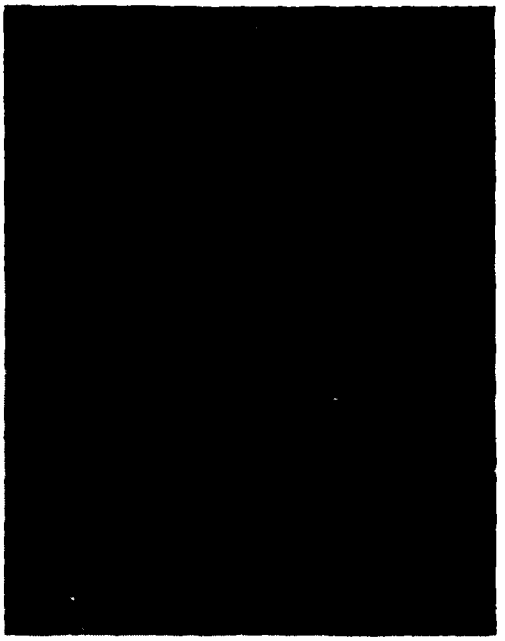


FIGURE 9
Cross-sectional TEM picture of an actual bipolar emitter's poly/single interface. The lattice fringes are continuous through the interface, indicating practically complete realignment of the poly.

devices. For large devices the base current was lower by a factor of two for the 800°C annealed ones. But, the I-V curves taken on small devices also indicated that their series resistance was substantially higher than of the devices given the 900 °C anneal. By a recently developed technique [19] it was determined that this was indeed due to emitter and not to base resistance. This resistance did not cause significant harm to the base transit times, extrapolated from F_t measurements, which are characteristic of low level current injection behavior. [20] However, when impact on switching performance was addressed in ECL and NTL ring-oscillators, the problems with the higher emitter resistance became obvious. At low power in the NTL circuits the smaller extrinsic base-collector capacitance (shallower junction) for the 800 °C annealed transistors allowed faster operation, 230ps delay vs 300ps. But at current levels typical of real circuit operations the delay times reversed, 135ps vs 165ps in favor of the 900°C

annealed ones, due to their lower emitter resistance. Similar results were obtained for ECL circuits.

3.5. Structural characterization

Clearly the interface between the poly and single crystal silicon must contain the clue for the difference in emitter resistance and the consequent performance difference. There were many studies done on such interfaces before. [21] However, our case was the first where such extensive electrical measurements were taken and correlated with a single processing detail relating to the polysilicon interface, and most significantly, also correlated with cross-sectional transmission electron microscopy (TEM) done on the very same emitters that have been electrically characterized. The result from the 900°C annealed device is shown on Fig. 9 in high magnification. The picture shows clear realignment of the poly, with almost total disappearance of the original interface. The reason for the low resistance was simply that for all practical purposes the interface vanished.

3.6. Conclusions

The presented picture is interesting from the point of view that it illustrates the care that must be exercised when one characterizes small devices. Based on projections from large devices, or even from small ones but at low current levels, one would have concluded that high gain is the most desirable property to pursue. The increase of emitter resistance accompanying the higher gain (low base current), showed up only in small devices, and its detrimental effects were only observed under realistic high-current injection testing. The implications for advanced bipolars are important. To achieve the highest speeds the best is if the interface is completely eliminated. A relatively "thick" single crystal emitter (which a regrown polysilicon contact becomes), coupled with a thin base is the goal to be attained, even if at low power levels this leads to an increased base current.

4. SELECTED TOPICS

The field of small devices is so vast that we could cover only a small fraction of it. There are many other areas worth of discussion as well. We would like express opinion on two topics, one well known, the other more exotic. The former is the quest to achieve good quality shallow junctions. We believe that much is to be gained in this field if one stays at very low, 550 to 800°C, temperatures and uses the property of many dopants to supersaturate and become active when silicon regrows from its amorphous state. [9,22,23] To mention a less well known topic, one has to be concerned with α-particle caused soft-errors for both small FET's [24] and bipolars. [25] As circuits respond ever faster, the time evolution of the initial current surge that follows a hit on one or more junctions is beginning to be of importance. It is almost universally assumed that the rise-time of such a current pulse is practically instantaneous, certainly of less than 1ps duration. [26] However, we believe that this time might be much longer, 10 to 30ps, due to the time it takes for the minority carriers to thermalize.

5. SUMMARY

We presented results and detailed several issues on the deeply-submicron world of silicon. Overall, we see nothing on the immediate horizon that would halt the ongoing progress of the technology.

6. ACKNOWLEDGEMENTS

The work presented in this article was a collaborative effort of many people in our laboratory. I'd like to draw special attention to references [5] and [15] from where many of the results and discussions given here were derived. During the preparation of this manuscript I received helpful advise from H. Stork and E. Ganin. H. Luhn is acknowledged for his help with the photographs.

REFERENCES

1. R.H. Dennard, F.H. Gaensslen, L. Kuhn, and H.N. Yu, IEEE IEDM Tech. Dig., p. 344 (1972).

2. G. Baccarani and M.R. Wordeman, Trans. Electron Devices. ED-30 1295 (1983).

3. J. P. Leburton and G. Dorda, Solid-State Elec., 26, 611 (1983).

4. J. G. Ruch, Trans. Electron Devices. ED-19 652 (1972).

5. G. A. Sai-Halasz, M. R. Wordeman, D. P. Kern, E. Ganin, S. Rishton, D. S. Zicherman, H. Schmid, M. R. Polcari, H. Y. Ng, P. J. Restle, T. H. P. Chang, and R. H. Dennard, in press

6. G. Baccarani, M.R. Wordeman, and R.H. Dennard, IEEE Trans. on Electron Dev., ED-31, 452 (1984).

7. M. R. Wordeman, A. M. Schweighart, R. H. Dennard, G. A. Sai-Halasz, and W. W. Molzen, IEEE Trans. Electron Devices ED-32, 2214 (1985).

8. S.A. Rishton, H. Schmid, D.P. Kern, H.E. Luhn, T.H.P. Chang, G.A. Sai-Halasz, M.R. Wordeman, E. Ganin, M. Polcari, Proceedings of the 31st Int'l. Symp. on Electron, Ion, and Photon Beams, Woodland Hills, CA, May 1987.

9. G. A. Sai-Halasz and H. B. Harrison, IEEE Electron Device Lett. EDL-7, 534 (1986).

10. P. M. Solomon, Proceedings of the IEEE, 70, 489 (1982).

11. G. G. Shahidi, D. A. Antoniadis, and H. I. Smith, IEEE IEDM Tech. Digest, p. 824 (1986).

12. S. Horiguchi, T. Kobayashi, M. Miyaki, M. Ode, and K. Kiuchi IEEE IEDM Tech. Digest, p 761 (1985)

13. E. M. Buturla, P. E. Cottrell, B. M. Grossman, and K. A. Salsburg, IBM, J. Res. Develop. 25, 218 (1981)

14. D. D. Tang and P. M. Solomon, IEEE H. Solid-Satate Circuits, SC-14, 679 (1979)

15. J. M. C. Stork, E. Ganin, J. D. Cressler, G. L. Patton, and G. A. Sai-Halasz, IBM, J. Res. Develop. Nov. 1987

16. T. H. Ning and R. D. Isaac, IEEE Trans. Electron Devices, ED-27, 2051 (1980)

17. H. C. de Graaff and J. G. de Groot, IEEE Trans. Electron Devices, ED-26, 1771 (1979)

18. J. M. C. Stork, C. Y. Wong, and .M. Arienzo, Symp. VLSI Technol. Dig. Tech. Papers, p 54 (1985)

19. J. M. C. Stork and J. D. Cressler, Symp. VLSI Technol. Dig. Tech. Papers, p 47 (1986)

20. See, for instance, I. Getreu, Modeling the Bipolar Transistor (Tektronix, Beaverton, Oregon, 1976) p. 172

21. See, for instance, G. L. Patton, J. C. Bravman, and J. D. Plummer, IEEE Trans. Electron Devices, ED-33, 1754 (1986)

22. E. Ganin, G. A. Sai-Halasz, and T. O. Sedgwick, Mat. Res. Soc. Symp. Anaheim, April 1987, and E. Ganin, D. L. Harame, and G. A. Sai-Halasz in press.

23. H. B. Harrison, Y. H. Li, G. A. Sai-Halasz, and S. S. Iyer, Mat. Res. Soc. Symp. 71, 223 (1986)

24. G. A. Sai-Halasz, M.R. Wordeman, and R.H. Dennard, IEEE Trans. Electron Devices ED-29, 725 (1982)

25. G.A. Sai-Halasz and D. D. Tang, IEEE IEDM Tech. Digest, p. 344 (1983)

26. See for instance, J. Chern, J. Seitchik, and P Yang, IEEE IEDM Tech. Digest, p. 538 (1986)

Session A1.2

Ultrasmall MOS Devices

Chairman: F.M. Klaassen

Monday, September 14, 1987

Half Micrometer N-MOS Technology Using X-Ray Lithography

V. Lauer, F.Bauer+, J. Korec++, H.-L. Huber*, P. Balk

Institute of Semiconductor Electronics
Aachen Technical University, D-5100 Aachen, FRG
* Fraunhofer-Institut für Mikrostrukturtechnik
D-1000 Berlin, FRG

MOSFETs with effective channel lengths down to 0.3 μm have been realized using x-ray lithography. To determine process parameters for device optimization two dimensional process and device modeling was employed. In addition, ring oscillators with different numbers of stages were fabricated to evaluate the performance of this technology.

INTRODUCTION

Fabrication of VLSI circuits not only requires a convenient high throughput technology but also demands processes which are capable of great accuracy. Improvements in mask technology, stepper alignment and resist technique have made x-ray lithography to a tool which is able to meet these high demands of circuit fabrication /1/. The aim of this work is to demonstrate the performance of x-ray lithography by using it for the definition of polysilicon gates in the halfmicron range. To swow the viability of this technique we want to present results on MOSFETs and ringoscillators obtained from an NMOS technology scaled to submicron dimensions. Devices fabricated with this process showed excellent IV-characteristics and almost no short channel effects. High speed and low power dissipation is an additional feature of the developed technology.

DEVICE MODELING AND PROCESSING

Process (SUPREM) and device (MINIMOS) modeling was employed to extend a 1μm NMOS process to the submicron range. Two conflicting boundary conditions /2/ had to be considered: minimum short channel effects and maximum device stability. Implantation profiles were optimized for inverters consisting of enhancement mode transistors with L_{eff}= 0.5 μm and depletion mode devices with L_{eff}= 0.9 μm. The internal electrical field, which is responsible for avalanche breakdown, carrier injection in the gate oxide and for the degradation of threshold and transconductance in submicron devices, was reduced by fixing the power supply to 3V. In this manner it was possible to avoid specially tailored profiles like lightly doped (LDD)/3/ or double diffused drain (DI-LDD)/4/. All lithography steps except the definition of the gate level were carried out with a deep UV contact printing system (Carl Süss, MJB 3 HP UV300).

Active device regions were defined by locally oxidizing the p-doped substrate (12-17 Ωcm). The interdevice isolation was improved by an additional boron implant. The threshold voltage of the enhancement mode transistors was adjusted to 0.5 V by a double boron implantation ($2*10^{12}$ cm^{-2}, 30 keV and $1*10^{12}$ cm^{-2}, 60 keV) through a sacrificial oxide of 25 nm. A

+ Now with Brown Boveri Corporation, Baden, Switzerland
++ Now with AEG Forschungslabor, Frankfurt, FRG

This work has been supported by the Bundesministerium für Forschung und Technologie

deep channel arsenic implant ($7.2*10^{11}$ cm^{-2}, 170 keV) was used for depletion load devices to obtain symmetrical inverter characteristics and to improve long time stability. Amorphous silicon (250 nm) was deposited in a low pressure process on the 15 nm gate oxide grown in dry oxygen at 1000°C. Per chip six transistors of each type with channel lengths varying from 0.8 to 1.8 µm for depletion mode and 0.5 to 1.2 µm for enhancement mode devices were structured by x-ray lithography and reactive ion etching by means of an SF_6 process. Channel width of all devices was 10 µm. For comparison an additional set of devices was prepared using e-beam direct writing for this step. After the deposition of a CVD passivation oxide the selfaligned arsenic source/drain implant (80 keV, 10^{16} cm^{-2}) was activated at 900°C resulting in a junction depth of about 0.2 µm. The Ti/Al (100/300 nm) metallization was structured by lift off.

To clarify the sensitivity of the described process to deviations in the technological parameters additional numerical simulations were performed. These demonstrated that a variation of channel length or oxide thickness up to ± 10% does not perturb the inverter characteristics considerably.

RESULTS

The ability of the used technology for manufacturing device structures of predetermined dimensions is demonstrated in fig.1. This figure shows the gate region of a 0.5 µm enhancement mode MOSFET. Excellent characteristics were obtained from the optimized devices (L_{eff}= 0.5 µm). Fig.2 shows the subthreshold current of an enhancement transistor with L_{eff}= 0.5 µm and W= 10 µm for different drain voltages. It may be seen that the curves are essentially insensitive to this parameter. This result indicates negligible short channel effects. The subthreshold slope of this device is 80 mV/decade.

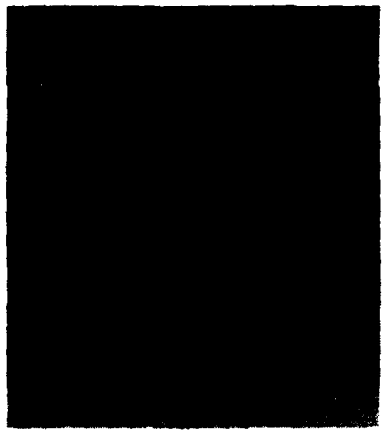


FIGURE 1

Gate region of MOSFET with L_{eff}= 0.3µm

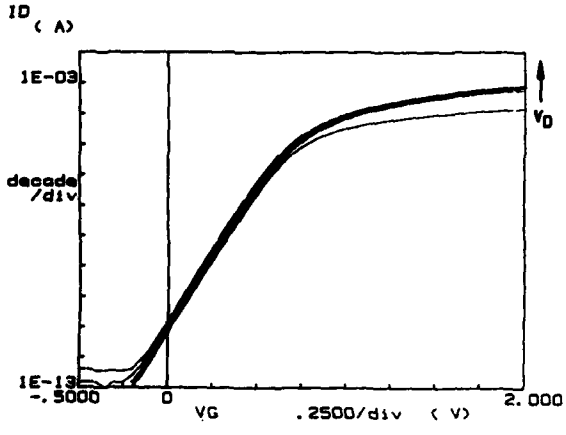


FIGURE 2

Subthreshold behavior of enhancement transistor with L_{eff}= 0.5 µm, W= 10µm
(0.1 V ≤ V_D ≤ 2.6 V, V_{SS}= 0 V)

As mentioned before, the shortest geometrical channel length used in this study was 0.5 µm. The corresponding IV-curves (fig.3) demonstrate that even these devices exhibit acceptable current voltage characteristics. Typical values for subdiffusion of ΔL= 0.19 µm and an overall series resistance of 24 Ω were determined using a method after Peng et al./5/. For this reason the shortest devices had an effective channel length of about 0.3 µm. Carrier mobility was measured with an approach proposed by Hao et

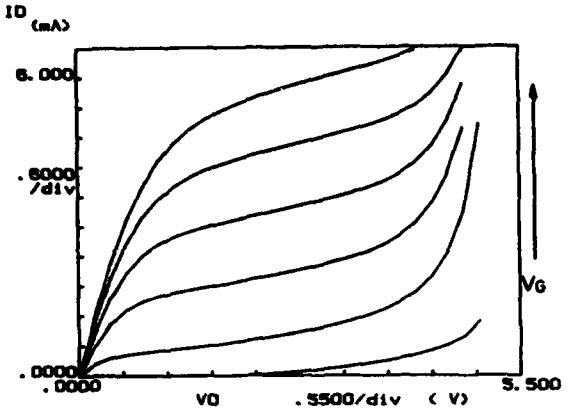


FIGURE 3

I_D-U_D characteristic of n-channel enhancement transistor (L_{eff}= 0.3μm)
(0 V ≤ V_G ≤ 5 V, V_{SS}= 0 V)

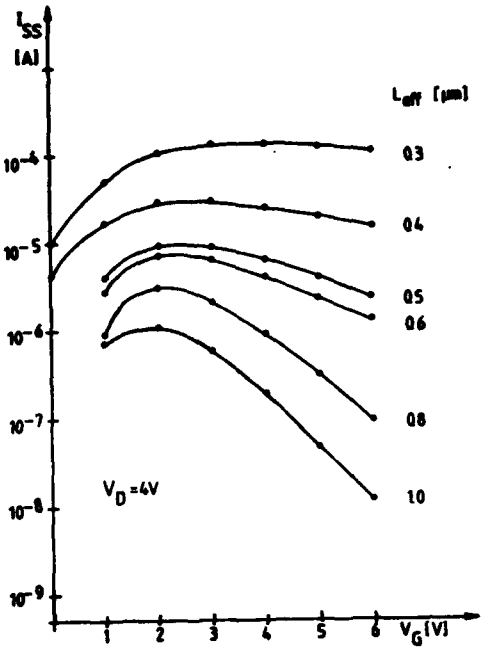


FIGURE 4

Substrate current of enhancement transistors with different channel lengths

al./6/. Values of μ_D= 530 cm^2/Vs for the depletion mode (L_{eff}= 0.9 μm) and μ_E= 425 cm^2/Vs for enhancement mode (L_{eff}= 0.5 μm) transistors are characteristic for devices in the submicron range /7/. The somewhat higher value for the depletion MOSFET is probably due to the buried channel. Very similar results were obtained on devices structured by e-beam direct writing.

Measured substrate currents for enhancement mode MOSFETs with different channel lengths are plotted in fig.4. Reduction of the effective channel length from 1.0 to 0.3 μm results in a rapid increase of substrate current over more than two decades. Also the shape of the I_{SS}-curves changes. The clearly visible maximum for the 'long-channel' devices at $V_G \approx V_D/2$ virtually disappears for transistors with channel length below 0.5 μm.

In order to get information about the dynamic potential of the described technology ringoscillators were designed with 0.5 μm effective channel length for the driver and 0.9 μm for the load transistor. Because it is well known that ringoscillators with a large number of stages tend to generate higher harmonics /8/ each chip contained 4 oscillators with different numbers of stages (101/51/25/13). A microphotograph of such a chip can be seen in fig. 5. The first results measured on oscillators which have been fabricated using x-ray lithography indicated a propagation delay time of 135 ps. A second set of devices was realized by means of e-beam direct writing with improved contact hole structuring and metallization. The output waveform for the 51-stage oscillator fabricated in this manner is shown in fig.6. This device has a typical inverter propagation delay time of 87 ps and a power delay product of 17.5 fJ at a supply voltage of 3V.



FIGURE 5

Microphotograph of ringoscillator chip (101/51/25/13 stages)



FIGURE 6

Output waveform of ringoscillator with 51 stages

CONCLUSIONS

Submicrometer MOSFETs with L_{eff} ranging from 1.6 to 0.3 μm have been realized using x-ray lithography for gate definition. Our results clearly show the ability of this technique for device fabrication. Excellent pattern definition and overlay accuracy are the two major features which have to be emphasized. The technology for the fabrication of these devices has been optimized employing numerical simulations. Short channel effects were almost completely suppressed by proper choise of the technological parameters. In addition, oscillators fabricated using this technology exhibit very high speed and low power dissipation.

REFERENCES:

/1/ A. Heuberger in Microcircuit Engineering, Ed. H.W. Lehmann, Ch. Bleiker (North-Holland, Amsterdam, 1986)

/2/ F. Bauer and P. Balk, Solid-St. Electronics, 29 (1986) 797

/3/ S. Ogura, C.F. Codella, N. Rovedo, J.F. Sheppard, J. Riseman, IEDM Techn. Digest (1982) 781

/4/ E. Takeda, H. Kume, Y. Nakagome, T. Makino, A. Shimizu, S. Asai, IEEE Trans. Electron Devices, ED-30 (1983) 652

/5/ K.-L. Peng, S.-Y. Oh, M.A. Afromowitz, J.L. Moll, IEEE Electron Dev. Lett., EDL-5 (1984) 473

/6/ C. Hao, B. Cabon-Till, S. Cristoloveanu, G. Ghibaudo, Solid-St. Electronics, 28 (1985) 1025

/7/ L. Manchanda, IEEE Electron Dev. Lett., EDL-5 (1984) 470

/8/ N. Sasaki, IEEE Trans. Electron Devices, ED-29 (1982) 280

OFFSET DIFFUSED DRAIN TRANSISTORS FOR HALF-MICRON CMOS

P. H. Woerlee, C. A. H. Juffermans, H. Lifka, F. M. Oude Lansink, H. J. H. Merks-Eppingbroek
T. Poorter and A. J. Walker

Philips Research Laboratories, P O box 80000
5600JA Eindhoven, The Netherlands

Half-micron n- and p-channel transistors with an offset diffused drain structure have been fabricated. A high temperature process and offset implantation of the source-drain dopants was used to obtain graded doping profiles and an optimum effective channel length. Experimental results showed good device quality. The extrapolated lifetime for the n-channel device was 5 years at a power supply voltage of 4.5 V. Hot carrier degradation in p-channel devices cannot be neglected anymore.

1. INTRODUCTION

Short channel effects and reliability aspects give rise to conflicting demands on the design of half-micron MOS transistors. A compromise has to be made when the transistors are designed for a fixed standard operating voltage.

It is the aim of this paper to investigate the feasibility of a high temperature budget half-micron CMOS process. The high temperature budget is useful for the reduction of the hot carrier degradation phenomena. A special drain structure in which the source-drain implantations were done after spacer formation was used to obtain graded source-drain profiles with limited lateral diffusion under the gate. In this way devices can be designed with a strong reduction of the lateral electric fields and with reasonable short channel effects. Device simulations were used extensively to get insight into the processing demands and for fine tuning of the doping profiles. Simulations, fabrication and experimental results for n- and p-channel devices will be discussed.

2. SIMULATIONS

The transistors were designed for a CMOS process with N-type polycrystalline silicon gate material,a gate length of $0.5\mu m$, a threshold voltage of 0.8 V, a gate oxide thickness in the range between 10 and 13.5 nm and a power supply voltage of at least 3.7 V. Furthermore the off-current at $V_{ds} = 4V$ should be below 10 $pA/\mu m$. Series resistances should not degrade the current drive capability in the saturation region by more than 10 percent. Simulation tools used in the study were DDF a 2D process simulator, and CURRY [1] a 2D device simulator. Both programs are proprietary software of Philips. The n- and p-channel devices will be discussed separately.

2. 1 n-channel transistors

The reduction of the hot carrier phenomena is the main problem in the design of the transistors. Lightly-doped-drain (LDD) [2] or doubly-diffused-drain (DDD) [3] structures have been used for this purpose. A lightly doped N^- region is used to reduce the lateral electric field E_x. For an LDD structure an N^- dosis below $10^{13}cm^{-2}$ gives the strongest reduction of E_x. However, such a low N^- dose leads to large series resistances and device degradation can even be more serious than in conventional devices [4]. Typically an N^- dose around $4 \times 10^{13}cm^{-2}$ gives optimum results.

To obtain a better understanding of the phenomena which lead to the reduction of E_x for such a high N^- dose, a series of simulations have been done. In these an LDD type drain structure with a gaussian N^- profile was used. The depth and lateral spread could be varied independently. The N^- top concentration was $5 \times 10^{18}cm^{-3}$. The junction depth (0.1 to $0.4\mu m$) was found to be of little importance for the reduction of E_x. The lateral spread of the N^- doping profile is the dominating factor in the reduction of E_x.

A simple way to produce a highly graded N^- profile is the use of a phosphorus N^- implant followed by a high temperature drive step. For a power supply voltage above 3.7 V, simulations showed that a drive step of $1000°C$ for 30 minutes is needed. For small devices the lateral diffusion is too large. Hence the N^- implants should be done after spacer formation.

Such an offset diffused drain structure (ODD) allows the formation of highly graded junctions with optimum effective channel length. A similar drain structure has been proposed by Horiuchi and Yamachuchi [5]. Simulations of 0.5 μm gate length ODD devices showed excellent device characteristics and high current drive capability. Furthermore the off currents are below 1 $pA/\mu m$ at $V_{ds} = 4$ V .

2. 2 p-channel transistors

The design of the buried channel PMOST has to be adjusted to the high temperature budget needed for the n-channel devices. An important design parameter of a buried channel device is the junction depth of the compensating boron threshold implant. To reduce drain induced barrier lowering phenomena this channel junction depth should be between 0.05 and 0.1 μm for our devices. This can be realized for the present process by the use of a medium energy anti punch through (APT) implant. The channel doping is shown in figure 1 for a 250 keV As^+ APT implant. The channel junction is very shallow even after the high temperature drive step. Furthermore the processing window is reasonable .

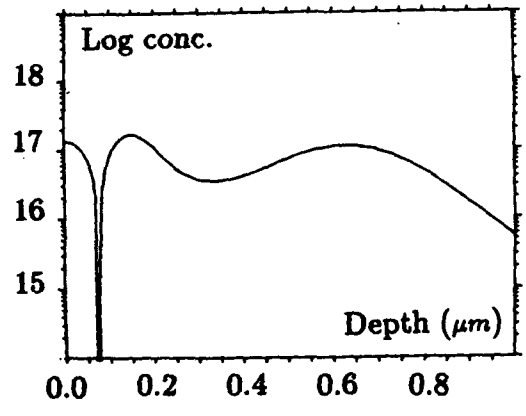


Fig. 1: Channel doping profile of the buried channel PMOST .

The source-drain P^+ implantation was done after the spacer formation, therefore the final high temperature step (925°C, 30 min) was used to form an offset graded drain. Simulations of the subthreshold characteristics of a 0.5μm device with gateoxide thickness of 12.5 nm are shown in figure 2. The subthreshold slope is 105 mV/dec at $V_{ds} = -3.7$ V.

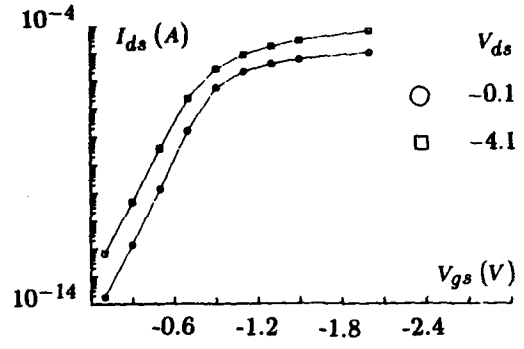


Fig. 2: Subthreshold characteristics of an 0.5μm gatelength PMOST .

3. FABRICATION PROCESS

The devices were fabricated using direct-write electron beam lithography. A retrograde n-well process using standard LOCOS field isolation was used. The gate oxides were grown in an oxygen/nitrogen mixture (10 percent oxygen by volume), the thickness was 10 nm and 13.5 nm for two different experiments. After surface and bulk implantations, 0.4 μm polycrystalline silicon was deposited and doped by arsenic implantation ($10^{16} cm^{-2}$). After an anisotropic gate etch and a sidewall oxidation step ,spacers of 0.1 μm width were formed. The source-drain formation was described in the previous section .

4. EXPERIMENTAL RESULTS

In figures 3 and 4 the I_{ds} - V_{ds} characteristics of an n-channel and a p-channel transistor with gate oxide thickness of 10 nm and a gatelength of 0.5μm are shown. The effective gatelength of the devices is 0.35μm. These graphs show a good saturation behaviour and a high breakdown voltage. The maximum transconductance in the saturation region is 200 $\mu S/\mu m$ and 100 $\mu S/\mu m$ for the n- and p-channel devices respectively.

In the linear region the current drive capability is limited by mobility degradation and series resistances. These effects remain of importance in the saturation region of the p-channel devices. The maximum mobility in the linear region is 500 cm^2/Vs and 125 cm^2/Vs for the n- and the p-channel transistors respectively. The series resistance for n-channel ODD's ($N^- = 7.5 \times 10^{13}$ cm^2) was extracted using the gate drive method [6]. Values of 300 $\Omega\mu m$ per side were found.

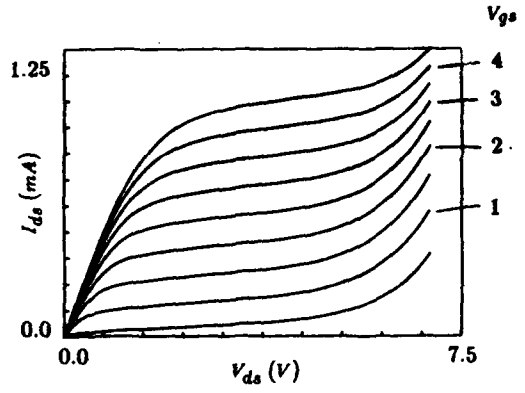


Fig. 3: $I_{ds}-V_{ds}$ characteristics of an ODD n-channel transistor. The gate length and channel width are 0.5 and 1.3 μm. N^- dosis is $7.5 \times 10^{13} cm^{-2}$.

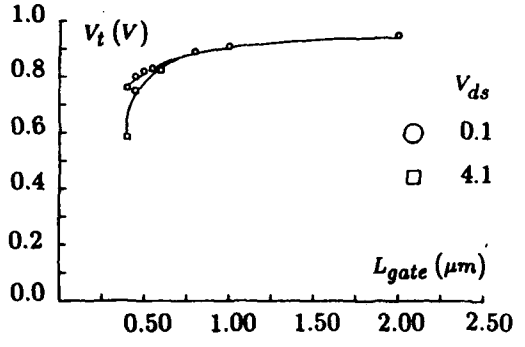


Fig. 5: Threshold voltage of n-channel devices as function of the gate length. $t_{ox} = 13.5\ nm$.

The dependence of the threshold voltage V_t on channel length at a drain bias of 0.1 and 4.1 V is shown in figures 5 and 6. Effects due to charge sharing and drain induced barrier lowering are relatively small for gate lengths larger than $0.4\mu m$ (n-channel) and $0.5\mu m$ (p-channel). For these channel lengths the leakage current at $V_{gs} = 0\ V$ and $V_{ds} = 4.1\ V$ was below 1 $pA/\mu m$. The body coefficient K of $0.5\mu m$ devices was 0.5 $V^{1/2}$. Narrow width effects on V_t and K were found to be negligibly small. This is caused by the relatively high channel doping levels of these small devices.

Finally the hot carrier degradation of the devices has been studied. The n-channel transistors with gate length of $0.5\mu m$ and $0.7\mu m$ were stressed at V_{ds} above 5.5 V and V_{gs} such that substrate current is maximum. As a lifetime criterion, a reduction of the device current of 10 percent at a fixed gate voltage and V_{ds} of 0.1 V was taken.

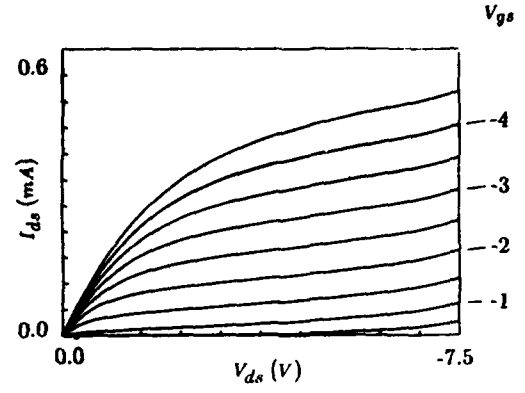


Fig. 4: $I_{ds}-V_{ds}$ characteristics of a p-channel transistor. Gate length and channel width are 0.5 and 1.2 μm .

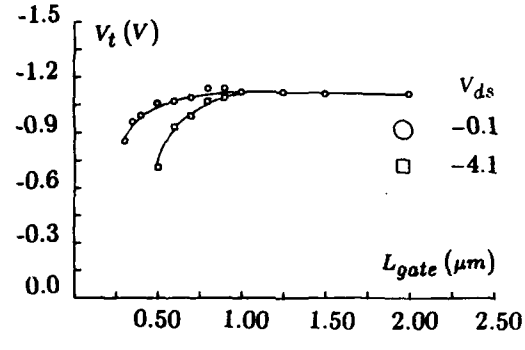


Fig. 6: Threshold voltage of the p-channel devices as function of the gate length. $t_{ox} = 10nm$.

The relative change in the current as function of the stress time is plotted in figure 7. Extrapolation of the device lifetime versus $1/V_{ds}$ to 5 years gave maximum allowable power supply voltages of 4.5 V and 5 V for the $0.5\mu m$ and $0.7\mu m$ devices .

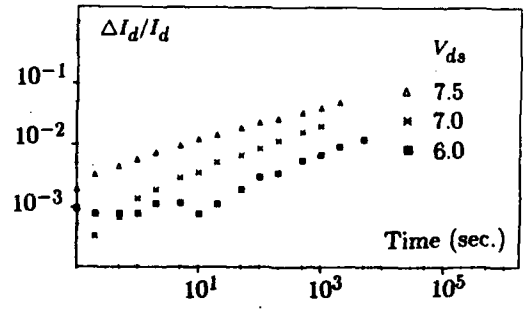


Fig. 7: Relative variation of I_{ds} at $V_{gs} = 3V$ as function of the stress time for an ODD n-channel device with $L_{gate} = 0.7\mu m$.

The p-channel devices show a different degradation behaviour. This is illustrated in figure 8 for a $0.5\mu m$ device. The stress conditions were $V_{gs} = -2V$ and $V_{ds} = -7V$. It can be seen in this figure that a very short stress time results in large changes in the subthreshold region. The leakage current increases by several orders of magnitude. Similar behaviour has recently been reported by Koyanagi et al [7], and it has been interpreted as hot electron induced punch through. Negative charge injected into the gate oxide causes a channel length shortening and this gives rise to an increase in punch through. Indeed in the larger devices the effects were found to be much smaller. Thus hot carrier effects are not negligible in half micron p-channel transistors and an LDD or ODD type drain structure has to be used .

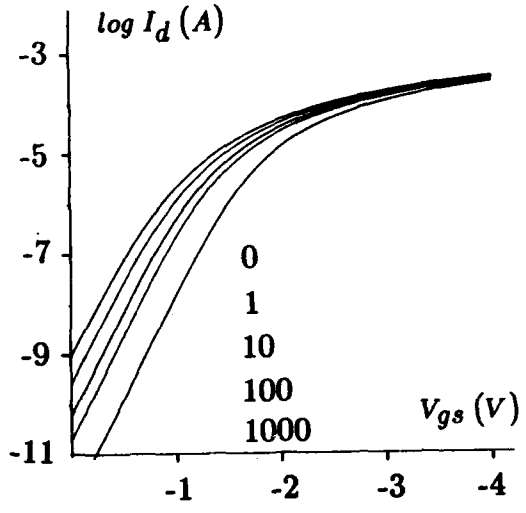


Fig. 8: Changes in the subthreshold behaviour at various stress times. Stress time in seconds is given in the figure .

5. CONCLUSIONS

The reduction of the hot carrier phenomena requires graded source-drain junction profiles when N^- dosis higher than $10^{13}cm^2$ are used. Hence a high temperature budget has to be used when hot carrier degradation phenomena need to be reduced. Excessive lateral diffusion under the gate can be avoided by offset implantation of the source-drain dopants. In this way the lateral electric field can be strongly reduced and an optimum effective channel length is obtained.

Half-micron devices with such an ODD drain structure were fabricated and showed excellent device characteristics, good performance and low series resistance. Furthermore, a 5 year lifetime at 4.5 V power supply voltage has been realised for the n-channel devices. The p-channel devices showed strong hot carrier degradation and need further improvement.

References

[1] Polak S J, Den Heijer C, Schilders W H A and Markowich P, Int. J. for Num. Meth. and Eng. ,vol 24, pp 763-838 (1987).

[2] Ogura S, Tsang P J, Walker W W,Critchlow D L and Shepard J F, IEEE Trans. Electron Devices ED-37, 1359 (1980).

[3] Hagiwara T, Yamaguchi K and Asai S, Symp. VLSI Technol. , Oiso, Dig. Tech. Pap. 46 (1982).

[4] Hsu F C and Grinolds H R, IEDM Tech. Dig. , pp 742-745 (1983).

[5] Horiuchi M and Yamaguchi K, Solid St. Electron. 28, 465 (1985).

[6] Terada K and Muta H, Japan. J. Appl. Phys. , 18, 953 (1979).

[7] Koyanagi M, Lewis A G, Zhu J, Martin R A, Huang T Y and Chen J Y, IEDM Tech. Dig. , pp 722-725 1986.

0.5 μm CMOS DEVICE DESIGN AND CHARACTERIZATION

H. I. Hanafi, M. R. Wordeman, L. K. Wang, Y. Taur, J. Y. C. Sun, R. H. Dennard, D. S. Zicherman, M. D. Rodriguez

IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598
And
N. Haddad, A. Edenfeld and M. Polavarapu
IBM FSD, Manassas, VA. U.S.A.

Abstract

The design and characterization of a high performance 0.5 μm channel CMOS is described. The design features thin epi with retrograded n-well, an n^+ polysilicon gate electrode, 12.5 nm gate oxide, shallow source/drain diffusions, and thin self-aligned titanium silicides. To control channel hot electron degradation effects in the NFET device with 3.3V power supply, different S/D junctions with graded profiles are investigated. The n-well doping profile is adjusted to provide adequate short channel threshold control and punch-through immunity in the buried channel PFET. In this paper, measured device characteristics will be discussed. Stage delays of unloaded inverter ring oscillators down to 90 pS are presented. Circuit performance sensitivities to a variety of parameters such as channel length, power supply and series resistance are also shown.

1. Introduction

Continuous scaling of CMOS technology is needed in order to achieve the required high speed and high packing density for VLSI. The technology described in this paper is aimed at improving performance through scaling down channel length by about 2x compared to a 1 μm CMOS design [1]. However, for TTL compatibility, the power supply is scaled down by only 1.5x to 3.3V. This increases the electric field intensity by about 1.3x in the 0.5 μm technology. To control channel hot electron degradation effects in the NFET device, S/D junctions with graded profile have to be achieved. To accomplish this, both As/P double-diffused and low dose As only junctions are investigated and discussed in Section 4. For process simplicity and higher hole mobility, an n^+ polysilicon gate is used for both the n- and p- channel devices. N-well doping profile is adjusted to provide adequate short channel threshold control and punch-through immunity in the buried channel PFET. Device characteristics and ring oscillator performance data are presented and discussed in sections 3 and 4 respectively.

2. Fabrication Process and Device Structure

Process steps of the CMOS technology described in reference [3] have been used to fabricate experimental devices and basic circuits. The n- and p- channel devices are designed with nominal channel length of 0.4μm and 0.5 μm respectively. An n^+ poly gate on 12.5 nm gate oxide and shallow source/drain junctions are used for both FETs (Fig. 1). To control channel hot electron degradation effects in the NFET device, two types of source/drain junctions were fabricated: As/P double-diffused and low dose ($< 5E15/cm^2$) arsenic-only junctions. The n- channel FET is built in a 0.75 Ω cm p-type epi-layer and the p-channel FET in a retrograded n-well with $7E16/cm^3$ surface concentration. A common BF_2 channel implant is used to adjust the n-channel and p-channel thresholds to +0.7 and -0.7V, respectively. Symmetry of thresholds about zero guarantees maximum noise immunity for logic circuits. To maintain proper threshold control in the short buried-channel PMOS, shallow channel doping profile and S/D junctions are required. A 0.2 μm p^+ junction is achieved using a silicon pre-implant to amorphize the

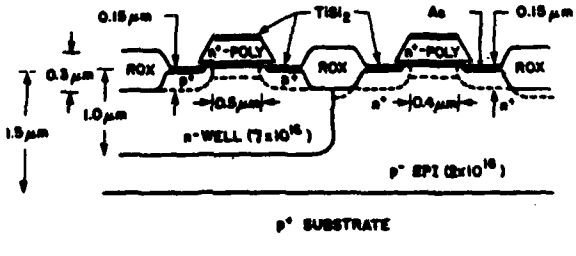


FIG. 1. 0.5um CMOS device cross sections.

substrate and avoid channeling of the subsequent boron ion implant [4]. The implants are followed by a short thermal cycle. Self-aligned $TiSi_2$ is formed over poly and n^+/p^+ diffusions with a sputtered titanium film [6]. A sheet resistance of 4-5 Ω/□ is obtained and the contact resistance contributions (between $TiSi_2$ and *Si*) to the device series resistance is less than 300 Ω μm [2,7] for both the n^+ and p^+ source/drain regions.

3. Device Characteristics

As mentioned in Section 2, graded S/D junctions in the NFET device are required in this technology to control channel hot electron degradation effects. To accomplish this, both As/P double-diffused and low dose arsenic only junctions were investigated. It was found that junction depths shallower than 0.25 μm are difficult to achieve using As/P double-diffused junctions due to an enhanced transient phosphorous diffusion. The use of As/P junctions in this design results in NFET punch-through at 0.45 μm channel length (Fig. 2). For channels longer than 0.45 μm, the NFET device is adequately turned off at zero gate bias with 80 mV/decade sub-threshold slope. For the As-only junction case, the junction depth was between 0.15 μm and 0.2 μm.

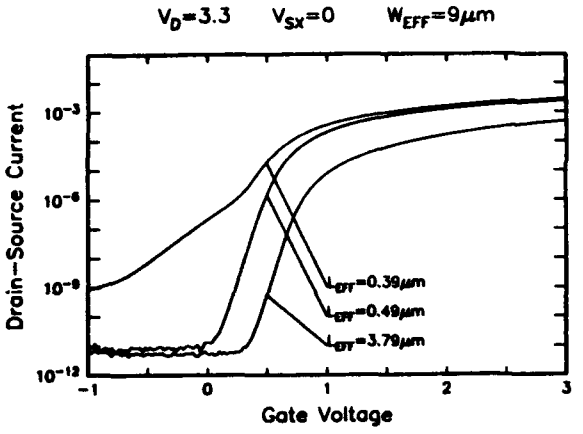


FIG. 2. Measured subthreshold characteristics of N-channel devices. (As/P S/D junctions)

For such shallow junctions, zero-degree angle source-drain implantation is used to avoid device asymmetry due to gate shadowing effects. Device punch-through is avoided in this case for channels longer than 0.3 μm as shown in Fig. 3. The dependence of the punch-through current on junction depth can be explained using the 2D device simulation shown in Fig. 4. In this figure, the sub-threshold drain current at V_{DS} = 3.3V is plotted against the gate voltage for a 0.3 μm channel length device with different source/drain junction depths. Note that as the source/drain junction depth increases, the sub-surface component of the drain current increases as well, which contributes to a higher off current. The measured short channel threshold fall-off characteristics of n-channel devices with As/P and As-only junctions is shown in Fig. 5. It is interesting to note that for both type of junctions, threshold fall-off is the same, i.e., independent of junction depths. This is also apparent from the device simulation results shown in Fig. 4. The reason for this is that, for high drain voltage and shallow S/D junctions, the channel depletion width is determined primarily by the junction curvature not its depth. This curvature is nearly independent of junction depth for very shallow junctions (< 0.3 μm).

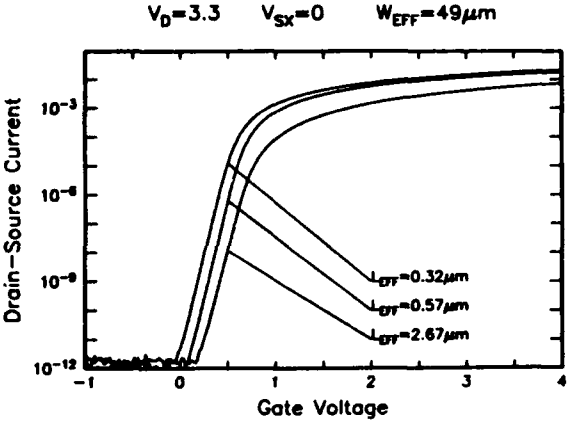


FIG. 3. Measured subthreshold characteristics of N-channel devices. (As S/D junctions)

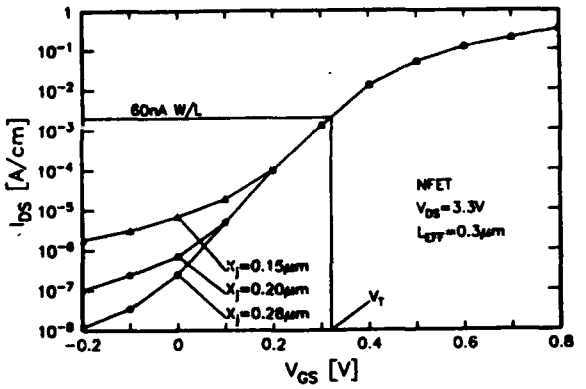


FIG. 4. Simulated NFET sub-threshold characteristics.

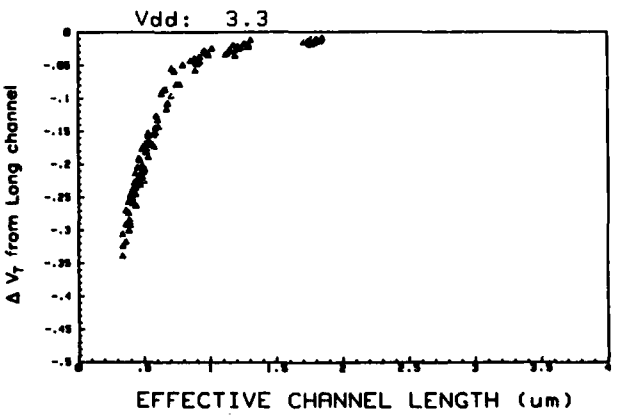


FIG. 5. NFET measured threshold voltage fall-off.

Both the As/P and the low arsenic dose result in graded S/D junction profiles which reduce the peak electric field at high drain biases. Hot electron reliability measurements show that the As gradient is such that the drain current of a 0.3 μm channel length device degrades by less than 20% when DC stressed with 3.7 volts power supply for 10 years as compared to a 3.9 volts power supply for the As/P case (Fig. 6). The NFET device series resistance (source and drain) measured using a channel length extraction technique [5] with small gate-bias is about 1250 Ω μm for the As/P case and 800 Ω μm for the As case. This allows a high saturation transconductance of 140 mS/mm for a nominal 0.4 μm channel length device with As S/D junctions (Fig. 7).

For process simplicity and high hole mobility, an n^+ polysilicon gate is used for both the n- and p-channel devices.

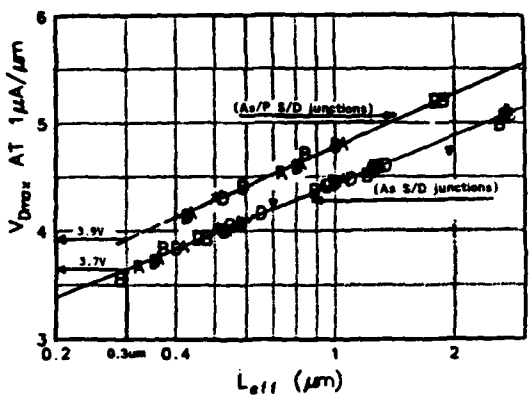


FIG. 6. NFET hot electron voltage limits.

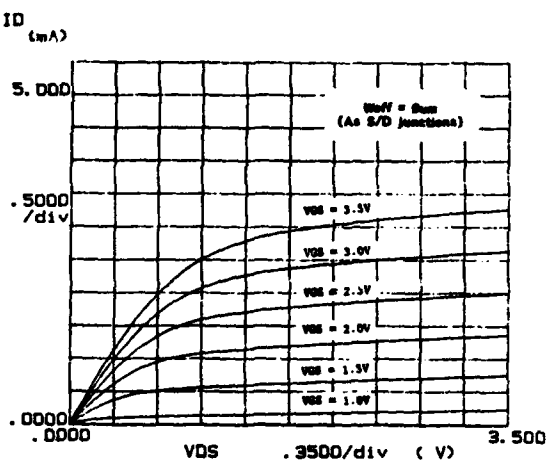


FIG. 7. I-V characteristics of NFET with 0.4um channel length.

This requires a counter doped channel implant in the p-channel device to lower the threshold voltage magnitude. The resulting buried p- channel device structure is known to have degraded short channel effects and turn-off behavior. However, by reducing the thermal cycle to keep the channel doping profile and S/D junctions shallow and with relatively high well doping concentration, short channel PFET devices with good characteristics are achievable. Figure 8 shows the measured sub-threshold characteristics of p- channel devices. Note that devices with 0.4 μm channel length are adequately turned off at zero gate bias and possess sub-threshold slopes of 95 mV/decade.

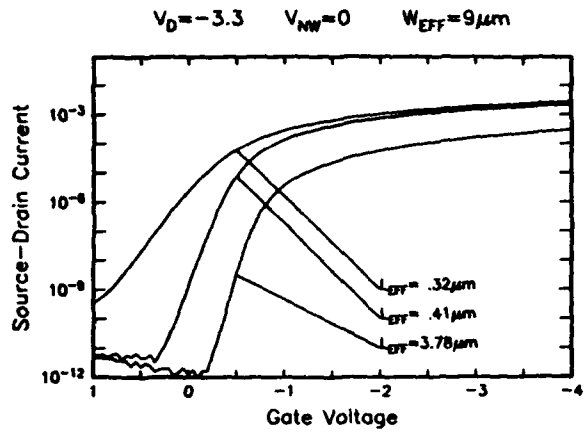


FIG. 8. Measured subthreshold characteristics of P-channel devices.

The short channel threshold voltage fall off of p- channel devices is shown in Fig. 9. These characteristics are similar to the NFET ones shown in Fig. 5 except that they are displaced by a 0.1 μm increase in channel length. The measured PFET device series resistance is 2000 Ω μm. (source and drain) and the 0.5 μm long device exhibits saturation transconductance of 85 mS/mm as shown in Fig. 10.

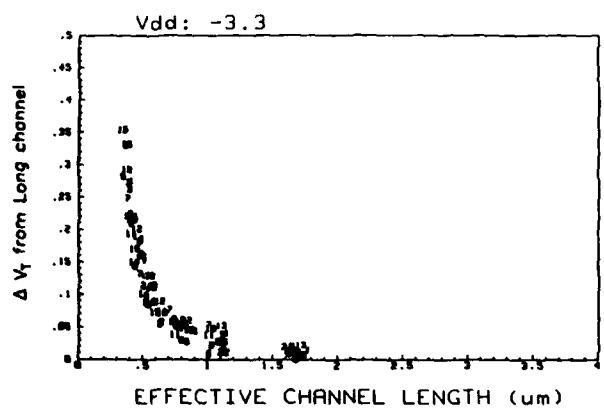


FIG. 9. PFET measured threshold voltage fall-off.

4. Circuit Performance

The dependance of circuit performance on power supply voltage measured on ring oscillators is shown in Fig. 11. These ring oscillators consist of 31 stages of unloaded CMOS inverters with fan-in and fan-out of 1. They are laid out with 1 μm ground rules. The minimum gate delay measured at

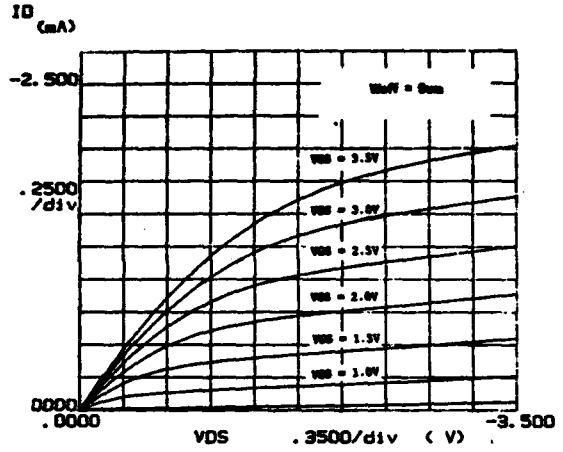


FIG. 10. I–V characteristics of PFET with 0.5um channel length.

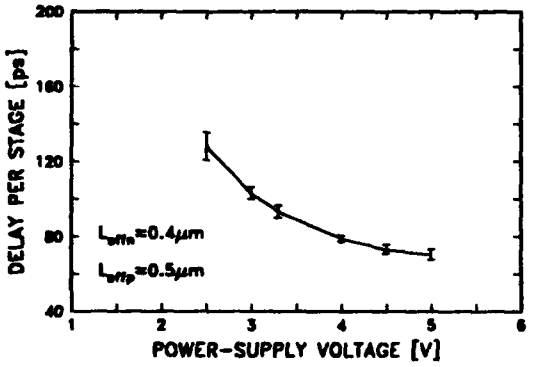


FIG. 11. Measured delay of performance ring oscillator.

3.3V is 90 pS. Figure 12 shows the performance of loaded ring oscillators with 15 stages having fan-in and fan-out of 3 in a NAND configuration with 0.2 μm loading capacitance (C_w) to simulate wiring capacitance. Note that the ring oscillators with NFET devices having As-only S/D junctions are about 15% faster than those with As/P junctions. This is mainly due to additional device series resistance and junction and overlap capacitance in the As/P case than the As case. At 0.45 μm channel length, 9 μm channel width, and V_{DD} = 3.3V, the average stage delay is 660 pS for the As junctions case.

5. Conclusion

The design, fabrication and characterization of a high performance 0.5 μm channel CMOS have been described. Experimental results for n- and p- channel device punch-through and short channel threshold fall off have been presented. Two NFET device designs were compared which showed a trade-off between device performance and reliability. Stage delays of unloaded inverter ring oscillators were measured down to 90 pS. Circuit performance sensitivities to channel length, power supply and series resistance were

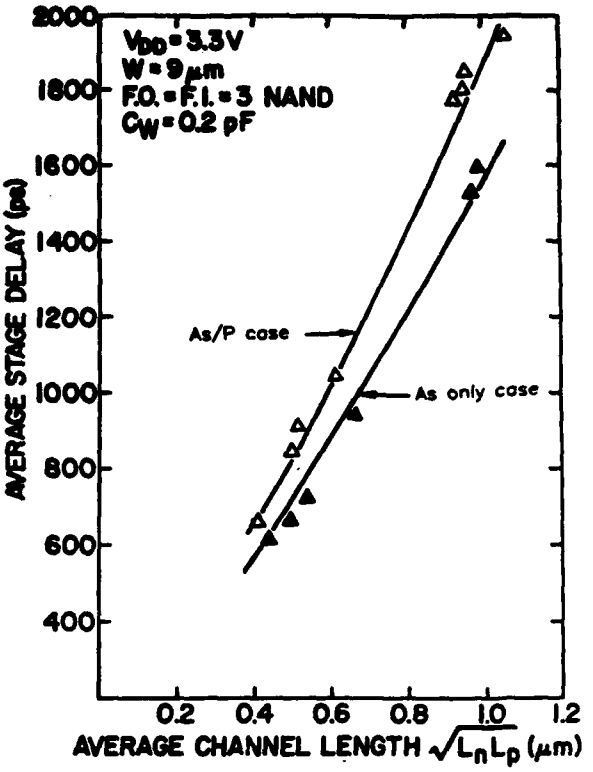


FIG. 12. Delay of loaded CMOS ring oscillators.

shown.

Acknowledgements

The authors would like to thank the personnel of the Yorktown and Manassas E-beam lithography groups, IBM Manassas FSD VLSI pilot line and the silicon facility at Yorktown Heights for their contributions to this work.

References

1. Y. Taur, et al., IEEE Trans. Elect. Dev., Vol. ED-32, p. 203, 1985.
2. H. I. Hanafi, 1986 ESSDERC Conferance Abstracts, P.85.
3. L. K. Wang, et al., Proc. 1986 Symp. VLSI Tech., P.13 (1986).
4. L. K. Wang, et al., presented at the Electrochem. Soc. Meeting, New Orleans, Louisiana, 1984.
5. K. Terada and H. Muta, Jpn. J. Appl. Phys., v.8, 953 (1979).
6. D. Moy, et al., Extended Abstract of Electrochem. Soc., Spring meeting, 1987, Vol. 87-1.
7. Y. Taur, et al., IEEE Trans. Elect. Dev., Vol. ED-34, p. 575, 1987.

Session B1.2

Monte Carlo Device Simulation I

Chairman: B. Riccò

Monday, September 14, 1987

MONTE CARLO SIMULATION OF SEMICONDUCTOR DEVICES: A CRITICAL REVIEW

Paolo LUGLI and Carlo JACOBONI

Dipartimento di Fisica e Centro Interuniversitario di Struttura della Materia dell'Universita' di Modena, Via Campi 213/A, 41100 Modena, Italy
Via Campi 213/A, 41100 Modena, Italia

We present a critical discussion of the Monte Carlo simulation as applied to semiconductor device modelling. The advantages and limitations of such approach are discussed and compared with more traditional simulators. Critical points are pointed out and analyzed. A variety of applications to different structures is then outlined.

1. INTRODUCTION

The Monte Carlo technique [1] is a fairly new tool in the area of device modeling, traditionally dominated by simulators based on drift-diffusion or on balance-equation models (for an overview of the subject see Ref. 2). Depending on the complexity of the simulated device, two main approches are generally followed. In the first one, a one particle Monte Carlo (OPMC) simulation is performed on a given fixed potential previously determined. In the second one, the traditional Monte Carlo simulation is performed for many particles in parallel (EMC) and coupled to Poisson's equation in order to obtain the self-consistent potential consistent with the charge distribution given directly by the Monte Carlo procedure.

Since no a-priori assumptions are needed on the form of the real and k-space carrier distributions, a Monte Carlo simulator is the only reliable tool for the investigation of those physical phenomena that critically depend on the shape of the distribution, or on the details of its tail (such as electron injection over potential barriers). Furthermore, the Monte Carlo technique allows us to focus on particular physical mechanisms that might be of importance on the device performance (for example, intercarrier scattering, impact ionization, generation-recombination, etc.). The prices one has to pay are a very time-consuming algorithm, and the requirement of a complete knowledge of the physical system under investigation. Often many assumptions have to be made in order to reduce the complexity of the model describing a given device.

The problem of the boundary conditions is of extreme importance in the M.C. simulation of semiconductor devices. Especially when the dimensions of the active region of a device are reduced below the micron limits, the overall performance of the device is largely controlled by contacts and boundaries.

Clearly, it will not be possible to exhaust the complexity of the M.C. simulation of devices in such a short review. Sec. 2 will present a short description of a generic M.C. simulator, followed by an overview at today's state of the art, outlining the area of applicability of the M.C. method, its advantages and its drawbacks. Many of the devices that have been modeled using a Monte Carlo method constitute very complicated systems, where high electron densities, sharp doping profiles, size quantization and heterojunctions between different materials are present, thus requiring very sophisticated algorithms. Sec.4 will deal with special features such as Pauli exculsion principle, tunneling and high energy effects. A complete overview of the Monte Carlo simulation of semiconductor devices will be presented in a forthcoming book [3].

2. DEVICE SIMULATION

In recent years the Ensemble Monte Carlo (EMC) has been widely used to study the properties of semiconductor devices. Particular emphasis has been lately attributed to submicron structures, because of their performances in switching and high frequency operations [4]. Once the basic physics involved in the transport of such devices is known,EMC simulation provides a formidable tool to determine their limits and characteristics and can be very helpful in modeling. Together

with the determination of the macroscopic properties of a device, EMC also gives a microscopic description of the local electric field, charge density,velocity distribution, etc.

A flow chart of a generic M.C. device simulation is shown in Fig.1. The basic steps are :

i) Charge assignement. A mesh of cells is set up and the charge of each particle is assigned to a particular mesh point. Since it is not possible to simulate all the electrons present in a real device, each simulated particle represents a cloud of electrons for the purpose of estimating currents, charge and field distributions. For all other purposes, each individual particle carries its elementary charge e. The doping charge is also added to the mesh according to its distribution.

ii) Potential solution. Poisson's equation is solved to determine the electrostatic potential at the mesh points. This is usually done in a very efficient way by using Fast Fourier Transform techniques. The electrostatic field field is then obtained from the potential with a finite-difference algorithm.

iii) Flights. Each particle,now treated as an individual electron, undergoes the standard MC sequence of scatterings and free flights, subject to the local field previously determined from the solution of Poisson's equation. The MC sequence is stopped at fixed times, when the field is adjusted following the steps described above. For OPMC techniques, the potential and field profiles are calculated at the beginning of the simulation, and only step iii) is performed. A detailed discussion of those issues is given in Ref. 5.

The description of the problem is completed by setting initial and boundary conditions. The initial conditions are not so important, since only the self-consistent steady-state result is usually retained. Boundary conditions are instead crucial, in particular in submicron devices, where contact properties drastically influence the whole behavior of the device.

Traditionally, device simulators have been based on drift diffusion (DD) or on balance equation (BE) models. Both of them are fast and reliable as long as a local description of the physical phenomena in the device is possible. That is, when the carriers can be described by a distribution characteristic of the given field present in every of the device. Such an assumption breaks down when the device dimension are small (typically below one micron), and high fields set up, leading to non-local phenomena. More specifically, when the field inside the device varies appreciably over length comparable with the electron mean free path, the electrons at a given position carries information about the field value at another position, and the trasport process becomes a non local phenomenon. The inclusion of the energy balance equation allows to incorporate some of these effects, at the cost of a much heavier computation. The Monte Carlo technique, which is inherently non local, lends itself very well to the simulation of

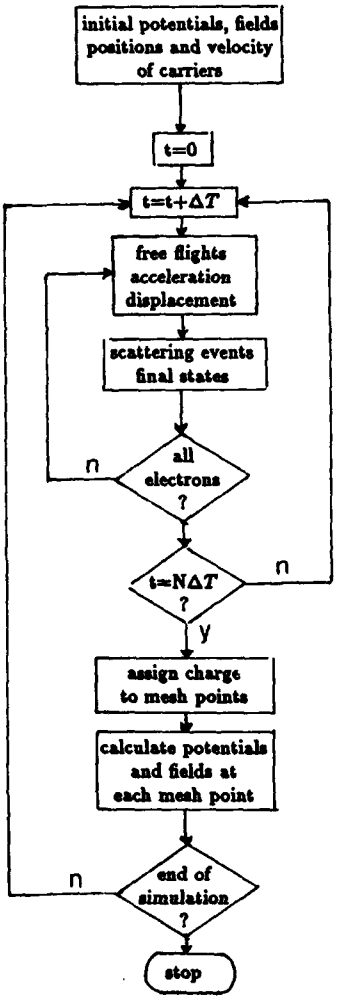


Fig. 1 Flow chart of a typical Monte Carlo device simulation

non stationary transport in dveices. The discrepancy between local models and the EMC have been clearly outlined by several authors [6,7,8].

A list of MC simulators present in the literature is shown in Tab. 1, together with the specific algorithm, and the respective references. The prototype, and hystorically the first self consistent MC program, was applied to a MESFET structure by Hockney and coworkers at Reading University [5,6]. The charge assignment and Poisson solver were based on algorithm deviced in the contest of plasma simulations. These algorithms have been so successful that they are adopted by most of the self consistent MC programs.

A large class of devices (MOSFET, bipolar transistors, HEMTS, etc) are characterized by areas with high doping (and free electron) concentrations, low electric field and retardind barriers. The direct simulation of electronic motion in these regions can be terribly time consuming. In contrast, more traditional simulators, such as (DD) schemes, can be applied to such a situation in a reliable and straightforward way. A hybrid method (HYMC) has been proposed [22,23] that combines the two techniques, by relying on the fast DD simulators for low field areas, and on the direct MC simulation where step gradients of the potential create the condition for hot carriers. Although excellent in principle, the hybrid technique (also called regional MC) requires a very accurate handling of the boundary conditions at the interface between the various region which ,in our opinion, has not yet been obtained.

Device	Simulation	Refs
MESFET	SSMC	6,8,9,10
HEMT	SSMC	11,12
MOSFET	ONMC,SSMC	13,14,15
HBT	HYMC	4
BIPOLAR	HYMC	16
PDB	EMC	17,18
THETA	SSMC	19,20
PBT	HYMC	21

Tab. 1 List of semiconductor devices simulated by Monte Carlo methods

3. SPECIAL FEATURES

In the following section, we focus on special aspects of the MC simulation that are non generally considered because of their difficulty, although they can be of great importance in the device performance.

3.1 Pauli Exclusion Principle

Electrons obey the Fermi-Dirac statistics and then must satisfy the Pauli Exclusion Principle. This means that not all the states are available because only one electron can reside in each state. This aspect is not very important in nondegenerate case and electrons are distributed in a large interval of states; in the degenerate case the problem becomes more conspicuous. In GaAs the electrons are degenerate for $T \leq 300K$ and for $n > 4.6 \times 10^{21} m^{-3}$. This is the case for many devices of interest. Degeneracy is equivalent to a many body interaction which reduces the phase space available for the electron final state in an induced transition.

If $p(\mathbf{k})$ and $p(\mathbf{k}')$ are the probabilities that respectively the initial and final state are occupied, the total rate of transition $P(\mathbf{k},\mathbf{k}')$ between two different states is given by $P(\mathbf{k},\mathbf{k}') = p(\mathbf{k}) \times S(\mathbf{k},\mathbf{k}') \times [1 - p(\mathbf{k}')]$. Normally a semiclassical Monte Carlo works with the approximatio $p(\mathbf{k}') = 0$ because all the states are considered available as final states. The inclusion of PEP is then essentially the inclusion of this term in the total scattering rate. In Ensemble Monte Carlo this is obtained very easily because the particle distribution is known step by step. The algorithm generating the distribution function is set up by multiplying the scattering probability by the correction factor $1 - p(\mathbf{k}')$; $p(\mathbf{k}')$ is determined self consistently and a rejection technique is used after selecting the final state without the correcting Pauli factor [24].

3.2 MOSFET simulation

One of the most important phenomenon in MOSFET's, and in general of submicron devices, is the heating of the channel electrons. As pointed out earlier, the EMC is most suitable technique in the presence of transient phenomena [25]. On the other side, the EMC poses considerable problems in MOSFET simulation, due to the very high doping of source and drain regions. A 2D OPMC has been proposed [15] that presents the following innovative features:

a) Using a piece-wise linear approximation of the elec-

tric field, the carrier motion equation are integrated analytically within each sector of the space grid thus avoiding time consuming numerical procedures.

b) All the boundaries of the simulation domain, of course except that along the $Si-SiO_2$ interface, have been fixed within low field device regions so that electrons can be injected with safe initial conditions (thermal distribution) and get out only after exhausting all important physical effects. This choice overcomes the questionable use of the boundaries posed by the HYMC.

c) An original, efficient algorithm has been implemented to calculate the appropriate duration of the free flights (depending on the actual carrier status). The method which is based on a space dependent definition of the scattering rate [26], leads to a drastic reduction in the number of selfscatterings thus allowing large saving in computation time (more than one order of magnitude compared with the conventional approaches).

d) The low injection efficiency of electrons from the n^+ source/ drain regions into the device channel has been overcome by means of the sample multiplication technique suggested in Ref.[27] to deal with rare carrier configurations. A multiplication technique of the same type been systematically used to reasonably populate the upper tails of the electron ED with an affordable number of total simulated particles. Good agreement with experimental values has been found, despite the simplicity of the physical model used to reproduce the MOS structure.

3.3 Shottky Diode Simulation

Metal-semiconductor contacts are of great importance in a number of semiconductor devices, whose applications range from high-speed logic to microvawes. As the dimensions of these devices reach the submicron limit, contacts become the limiting factor for the performance in the ballistic or quasi-ballistic mode of operation. In the same limit, the standard theory of Schottky barriers becomes less and less accurate.

A simulation of a 1-D metal$-n-n^+$ structure has been proposed [28], that presents an interesting system in that the device is never charge neutral, except under flat-band condition. This is due to the presence of a depletion or an accumulation region near the interface. Since the value of the electric field at the two boundaries $x=0$ and $x=w$ (see Fig.2) is related through Gauss' law to the net charge inside the device, it is necessary to allow the number of the electrons simulated to vary during the simulation, as a constant number of the electrons would give incorrect results. In this respect, a novel scheme, based on a cubic Hermite collocation method to solve the 1-D Poisson's equation has been introduced. The value of the electric field on the $x=0$ boundary is obtained without any loss of precision, and allows to obtain directly the value of the current through that interface. By carefully controlling the flux of the electrons in the two directions, it is possible to update constantly the numbers of electrons simulated, in accordance to the field distribution inside the device. In this way, the interface at $x=0$ is modelled as a perfectly injecting contact. The metal contact at $x=w$ acts as an absorbing boundary: electrons with energy sufficient to overcome the barrier are injected into the metal.

The tunneling probability for an electron with energy ε at a distance x from the interface is given by

$$T(\varepsilon) \cong \exp -2/\hbar \int_x^w dx 2m^*[qV(x)-\varepsilon]^{\frac{1}{2}} \qquad (1)$$

where m^* is the effective mass, V the potential seen by the electron, and $\hbar$ is Plank's constant. As a M.C. electron reaches the barrier, the tunneling probability is calculated from (1) using a parabolic least square interpolation of the potential $V(x)$ obtained from the solution of Poisson's equation. A random number is then used to decide whether the electron will tunnel or not. It is important to notice that no assumptions on the electron distribution function near the contact or on the shape of the potential barrier are needed.

The validity of the self-consistent algorithm has been verified by comparing the equilibrium potential distribution (reflecting the shape of the conduction band) obtained from the E.M.C. simulation with the result of an iterative process which combines the use of Fermi-Dirac statistics and Poisson's equation. At room temperature and moderate electron densities, current is provided solely by thermoionic processes. Higher densities cause a narrowing of the barrier, thus increasing the tunneling probability. The method indicates that significant tunnel currents start to appear for electron

concentrations above $5 \times 10^{18} cm^{-3}$. This indicates that the method proposed here might be suitable also for the simulation of ohmic contacts.

In summary, we have presented a critical overview of Monte Carlo simulations of semiconductor devices, outlining various approaches present in the literature, and areas of applicability

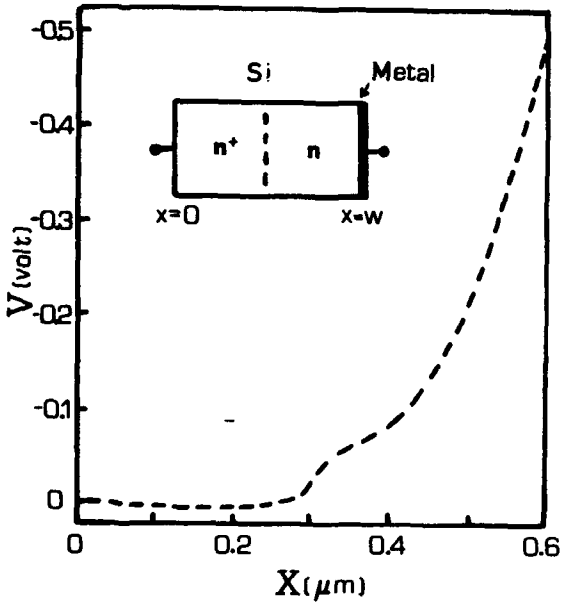


Fig. 2 Potential profile in the simulated Shottky diode.

ACKNOWLEDGEMENTS

This work is partially supported the Computer Centers of The University of Modena and by The National Research Council (CNR) under the Finalized Research Project Materials and Devices for Solid State Electronics.

REFERENCES

1. C. Jacoboni, and L. Reggiani, Rev. Mod. Phys. **55**, 645 (1983)
2. S. Selberherr, " Analysis and Simulation of Semiconductor Devices", Springer Verlag, Wien (1984)
3. C. Jacoboni, and P. Lugli, " Monte Carlo Simulation of Semiconductor Devices", to be published by Springer Verlag
4. R. Castagnè, in "High Speed Electronics", Eds. B. Källbäck and H. Beneking, Sringer Verlag, p.2 (1986)
5. R.W. Hockney, and J.W. Eastwood, "Computer Simulation using Particles", Mc Graw-Hill, N. York (1981)
6. R.W. Hockney, R.A. Warriner, and M. Reiser, Electron. Lett. **10**, 485 (1974)
7. W. R. Curtice, IEEE Trans. Electr. Dev. **ED 29** , 1942 (1982)
8. Yoshii, M. Tomizawa, and K. Yokoyama, IEEE Trans. Electron Dev. **ED30**, 1376 (1983)
9. Y. Awano, K. Tomizawa, and N. Hashizume, IEEE Trans. Electron Dev. **ED 31**, 448 (1984)
10. C. Moglestue, COMPEL **1**, 7 (1982)
11. M. Tomizawa, A Yoshii, and K. Yokoyama, IEEE Electron Dev. Lett. **EDL 6**, 332 (1985)
12. U. Ravaioli, and D.K. Ferry, IEEE Trans. Electr. Dev. **ED 33**, 677 (1986)
13. Y. Park, T. Tang, and D.H. Navon, IEEE Trans. Electr. Dev. **ED 30**, 1110 (1983)
14. K. Throngnumchai, K. Asada, and T. Sugano, IEEE Trans. Electron Dev., **ED-33**, 1005 (1986)
15. B. Riccò, E. Sangiorgi, F. Venturi, and P. Lugli, in Proc. Int. Electron Dev. Meeting (IEDM), Los Angeles, Ca, p. 559 (1986)
16. Y. Park, D. H. Navon, and T. Tang, IEEE Trans. Electron Dev., **ED-31**, 1724 (1984)
17. M.A. Littlejohn, R.J. Trew, J.R. Hauser, and J.M. Golio, J. Vac. Sci. Technolog. **B1**, 449 (1983)
18. T. Wang, K. Hess, and G.J. Iafrate, J. Appl. Phys. **59** , 2125 (1986)
19. F. Antonelli, and P. Lugli, submitted for publication
20. F. Antonelli, and P. Lugli, this book.
21. C. Hwang, D. H. Navon, and T. Tang, IEEE Trans. Electron Dev., **ED-34**, 154 (1987)
22. P. Nguyen, D. H. Navon, and T. Tang, IEEE Trans. Electron Dev., **ED-32**, 783 (1985)
23. S. Bandyopadhyay, M.E. Brown, C.M. Maziar, S. Datta, and M.S. Lundstrom, IEEE Trans. Electron Dev., **ED-34**, 392 (1987)
24. S. Bosi, and C. Jacoboni, J. Phys. C **9**, 315 (1976)
25. J. Zimmermann, and E. Constant, Solid State Electr **23** , 915 (1980)
26. E. Sangiorgi, B. Riccò, and F. Venturi, to be published in IEEE Trans. on CAD.
27. A. Phillips, and P.J. Price, Appl. Phys. Lett. **30**, 528 (1977)
28. U. Ravaioli, P. Lugli, M. Osman, and D. K. Ferry, IEEE Trans. Electr. Dev. **ED 32**, 2097 (1985)

A MONTE CARLO STUDY OF DIFFUSION COEFFICIENTS OF TWO-DIMENSIONAL ELECTRON GAS IN HEMT AlGaAs-GaAs STRUCTURES

J. Zimmermann, and Y. Wu

Centre Hyperfréquences et Semiconducteurs, UA CNRS n° 287
Bâtiment P3, Université de Lille-Flandres-Artois
59655 Villeneuve d'Ascq Cedex, France

This paper presents a Monte Carlo study of diffusion and noise in two-dimensional electron gas (2DEG) in heterojunctions. This is mainly achieved via the calculation of velocity fluctuations correlation functions of the 2DEG subjected to a driving field applied along the channel. It is found that at rather low fields when the carriers have a real two-dimensional motion, the parallel correlation functions show oscillations which we analyse in terms of scattering rates. This gives rise to 2DEG resonant noise spectra whose resonance frequency is very nearly proportional to the driving field strength.

As it is well known, one of the major advantages of Hot Electron Mobility Transistors (HEMT) comes from the fact that they present a good noise figure associated with a high gain at microwave frequencies, as compared to the normal MESFET |1|. There are at least four main sources of noise: diffusion noise, shot noise, G-R noise and 1/f noise. At high frequency, the only which prevails is diffusion noise. For these reasons, diffusion coefficients have been extensively studied in a recent past for a number of bulk materials: Si, GaAs, InP, GaInAs |2|-|6|.

THE MODEL

Basically, in any Monte Carlo model, two ensembles of data are needed: i) the energy bands in which the carriers move; ii) the various scattering probabilities.

Our simulations being aimed at obtaining valuable results for real structures, the following structure has been considered. From the GaAs SI-substrate, an undoped (π) GaAs buffer layer is grown, followed by an undoped $Al_xGa_{1-x}As$ (x=0.3) spacer and a Si-doped (N_d =1.3x10^{18} cm-3) AlGaAs (x=0.3) layer on which a metal gate is deposited. The widths are 1µm, 50 Å, and 625 Å respectively.

i)Energy bands. On the GaAs side of the heterojunction a potential well is formed, the shape of which is controlled by the electron charge trapped in it, which in turn is controlled by the gate bias voltage.
The subbands are obtained by a self-consistent solution of the Schrödinger and Poisson equations in which we assume an electron effective mass to remain valid |7|,|8| and an apparent activation energy of the dopant in the AlGaAs layer |9|.

ii) Scattering probabilities. Detailed accounts on the way in which 2DEG scattering probabilities can be evaluated can be found in a number of references |10|-|14|.

The conduction band scheme appears as conduction bands of a normal AlGaAs-GaAs system connected to a subsidiary Γ valley in GaAs where the subband structure is fully taken into account. At low fields, the carriers move essentially in the latter. Of course, we allow transitions between 2D states and 3D states to occur in which the localized electron wave-functions in the sub-Γ valley play an essential role. A full account of this part of the model can be found in ref.15.

SIMULATION RESULTS

We will now first show and illustrate velocity fluctuations correlation functions, since their Fourier Transforms gives the diffusion noise power spectra which in principle can be directly measured experimentally |16|. Figures 1a,b show parallel and transverse correlation functions at driving fields of 200 and 600 V/cm respectively. All these results were obtained with n =1.1x10^{12}

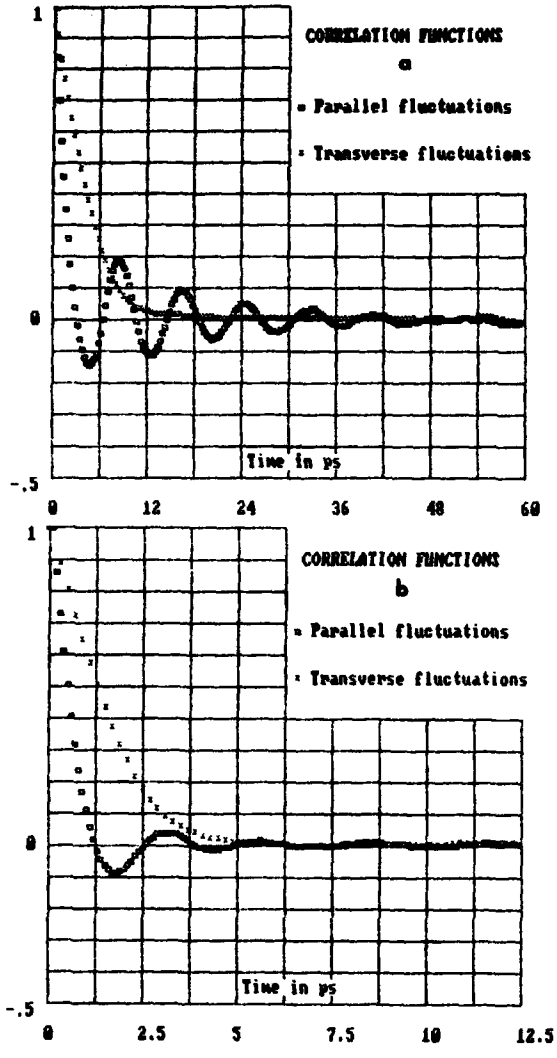


Fig. 1. In (a) parallel and transverse correlation functions computed at 200 V/cm for n_s =1.1x 10^{12} cm-2 and 77 K. In (b) the same are computed at 600 V/cm. The functions are reduced to 1 at t=0.

cm-2 at 77 K. One observes that while transverse correlation functions exhibit a nearly exponential time decay, the parallel correlation functions exhibit strong oscillations whose amplitudes also decay in time. We have noticed that the pseudo-period of these oscillations is inversely proportional to the electric field strength, and the oscillations disappear at vanishing fields and above 1000 V/cm. We have noticed that the amplitudes of these oscillations decrease as n_s is lower. This is illustrated in fig. 2 where the correlation functions calculated at 200 V/cm at 77 K with n_s =1x10^{11} cm-2 are displayed. The corresponding noise power spectra of these various cases are shown in figs. 3a,b and in fig. 4. In the next section we propose an explanation for this phenomenon.

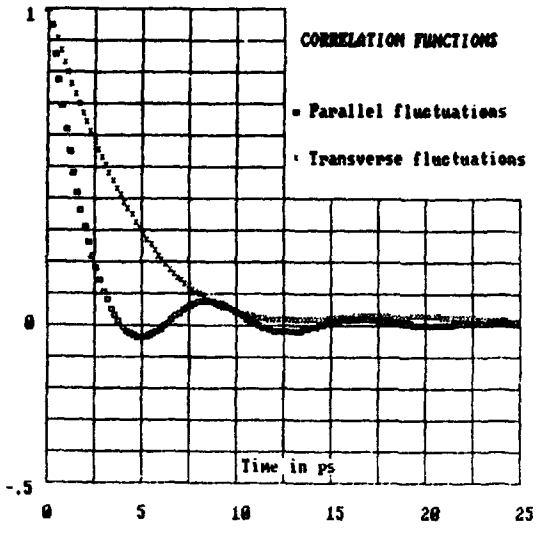


Fig 2. Parallel and transverse correlation functions calculated at 200 V/cm for n_s =1x10^{11} cm-2 and 77 K. The functions are reduced to 1 at t=0.

Mean-square displacements can also be evaluated in the simulations. An example is shown in fig.5 in a case where E= 600 V/cm and n =1.1x10^{12} cm-2 at 77 K. Linearity is clearly observed at times greater than 50 ps.

DISCUSSION and COMMENTS

The fact that a velocity correlation function may exhibit oscillations had been anticipated by Ferry and Barker |17| based on a simple Shockley model. In our case this can be understood in the following way. The various assumptions in the model are: i) in the quantum well polar optical phonon scattering dominates and is strongly anisotropic (small angle scatterings only are important), phonon absorption is always weak, while phonon emission becomes very important as soon as

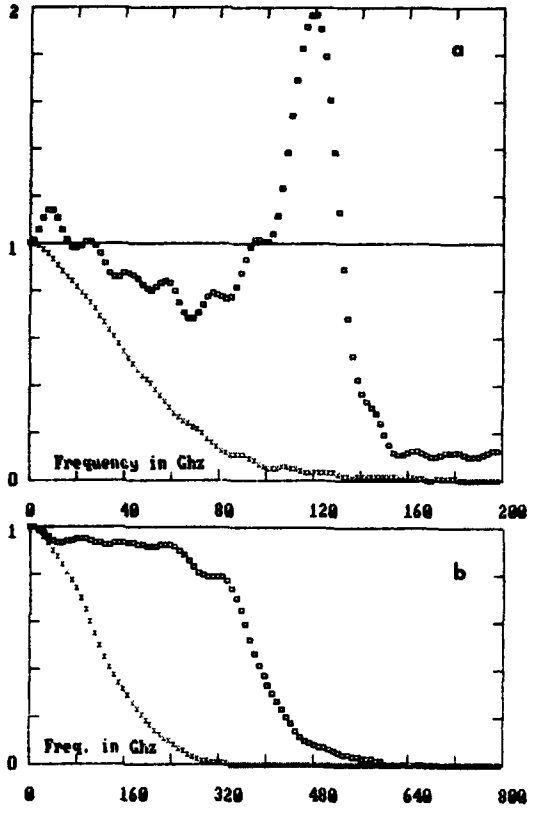


Fig. 3. Parallel (▫) and transverse (x) noise power spectra.(a): spectrum of fig 1a. (b): spectrum of fig. 1b. The spectra are reduced to 1 at zero frequency.

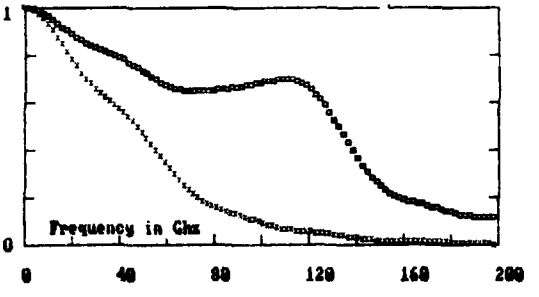


Fig 4. Parallel (▫) and transverse (x) noise power spectra of the correlation functions of fig. 2.

the electron energy reaches the energy of the phonon, namely 35.4 meV; piezo-electric scatterings exist but are elastic and anisotropic; elastic acoustic scatterings are radomizing (strongly isotropic) at a constant scattering rate |10|; impurity scattering (elastic) is anisotropic and rather weak due to the presence of the spacer

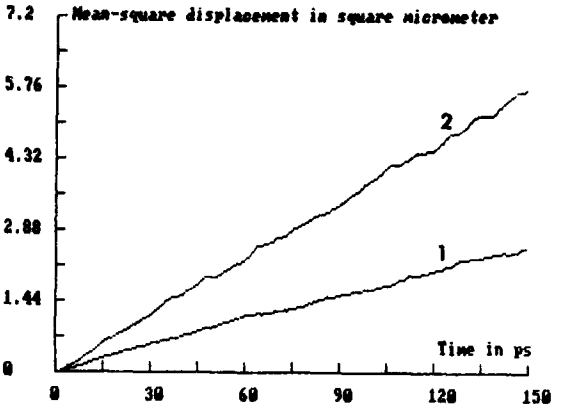


Fig. 5. Parallel (1) and transverse (2) mean-square displacements of a bunch of carriers computed at 600 V/cm for n_s =1.1x10^{12} cm-2 at 77 K.

layer. ii) At n =1.1x10^{12} cm-2 the separation between the ground and second subbands is 37 meV; at 1x10^{11} cm-2 it is 16 meV.

If we considered that polar-optical phonon (POP) scattering acted alone, then a carrier with a zero velocity at the beginning would be accelerated until its energy reaches POP energy, at which time it would be very likely to emit a phonon and have its energy sent back near the subband edge. The time needed for this process to occur is

$$T = 1/qE.(2m^*\hbar\omega_{pop})^{1/2}, \qquad (1)$$

in which case the velocity fluctuations correlation function would exhibit oscillations with the period T , velocity states separated by T would almost completely correlate, the correlation function would decay very slowly or not at all. In our case though, the additional collisions mainly due to isotropic acoustic phonon scattering randomize at a rather slow rate the electron velocity during the acceleration periods and thus have two main effects: the oscillations will be damped due to these extra collisions, the pseudo-period will be slightly shorter than T but the 1/E dependence should be preserved. This is exactly what we observed from the simulated correlation functions. We summarize the results in Table 1.

The analysis of the amplitudes of the oscillations is less easy. Figure 1a,b shows examples where the separation between the first two subbands

TABLE 1

Field	Period T	M.C. period
100 V/cm	16.4 ps	15.8 ps
200	8.2	8
400	4.1	3.6
600	2.7	2.2

is larger than the POP energy. In that case inter-subband transfer assisted scattering is very weak; the fact that the amplitudes, when the field is higher, are less important is a result of the increasing number of POP collisions which are small angle but non-zero angle collisions. Figure 2 shows a case where the separation between the first two subbands is much lower than the POP energy. Inter-subband transfer assisted scattering is much stronger than in fig. 1a, and as a result a significant portion of the electron kinetic energy is transformed into potential energy, thus reducing velocity fluctuations while preserving the periodicity of the motion which is a function of the field and POP energy only.

We also have analyzed these observations by comparing them with similar simulations of electrons in bulk-GaAs at 77 K. In that latter case no oscillations were observed in the range of fields of interest here. We believe this to be a result of the fact that elastic acoustic scattering rate increases with energy instead of remaining a constant as for two-dimensional electrons. Then the oscillatory nature of the electron motion induced by POP collisions would be probably lost.

Diffusion noise spectra are obtained from correlation functions via Fourier transforms |3|. The oscillatory nature of correlation functions is reflected in spectra through the presence of a resonance frequency equal to 1/T . This clearly appears in figs. 3a and 4, where at 200 V/cm resonance occurs near 120 GHz. In general as frequency increases, noise power at first decreases and then increases more or less sharply at the resonance frequency. Above the resonance, noise power decays down to zero. In contrast, in the direction transverse to the field axis, noise power spectra decay monotonously to zero.

As a conclusion, it is interesting to note that this kind of behaviour in the low field range at rather low temperature is typical of a two-dimensional electron system as it is represented by our model. In the bulk, in the same conditions, no resonance occurs unless the field is greater than the threshold field for intervalley transfer.

REFERENCES

1. A. Cappy et al., IEEE Electron Dev.Lett. EDL-6 (1985) 270.
2. R. Brunetti, and C. Jacoboni, Phys.Rev.B 29 (1984) 5739.
3. R. Fauquembergue, J. Zimmermann, A. Kaszynski, and E. Constant, J.Appl.Phys. 51 (1980) 1065.
4. G. Hill, P.N. Robson, and W. Fawcett, J.Appl. Phys. 50 (1979) 356.
5. B.R. Nag, S.R. Ahmed, and M. Deb Roy, Solid-State Electron. 30 (1987) 235.
6. J. Zimmermann, P. Lugli, and D.K. Ferry, J.Physique (France) 42 (suppl.nr. 10,coll. C7) (1981) 95.
7. The parameters concerning AlGaAs and AlGaAs-GaAs heterojunctions are taken from S. Adachi, J.Appl.Phys. 58 (1985) R12.
8. B. Vinter, Appl.Phys.Lett. 44 (1984) 307.
9. N. Chand, T. Henderson, J. Klem, W.T. Masselink, R. Fischer, Y.C. Chang, and H. Morkoç, Phys.Rev.B 30 (1984) 4481.
10. P.J. Price, Annals of Physics (New York) 133 (1981) 217.
11. P.J. Price, J.Vac.Sci.Technol. 19 (1981) 599.
12. J. Zimmermann, Y. Wu, and F. Perri, Physica 129 B (1985) 385.
13. K. Yokoyama, and K. Hess, Phys.Rev.B 33 (1986) 5595.
14. J. Lee, H.N. Spector, and V.K. Arora, J.Appl. Phys. 54 (1983) 6995.
15. J. Zimmermann, and Y. Wu, Rev.Phys.Appl. (France) to be published in 1987 (in print).
16. J. Zimmermann, S. Bonfils, Y. Leroy, and E. Constant, Appl.Phys.Lett. 30 (1977) 245.
17. D.K. Ferry, and J.R. Barker, J.Appl.Phys. 52 (1981) 818.

MONTE-CARLO SIMULATION OF CLASSICAL AND INVERTED MODFET STRUCTURES

FAUQUEMBERGUE R., PERNISEK M., THOBEL J.L., BOUREL P.

centre hyperfrequences et semiconducteurs, LA CNRS 287 Bat. P3

universite des sciences et techniques de LILLE

59655 VILLENEUVE D'ASCQ , FRANCE

INTRODUCTION

MODFET structures consist mainly of a heterojunction located between a larger band gap highly doped material and a smaller band gap undoped material. The carriers of the larger band gap material diffuse into the smaller band gap material where they confine near the heterojunction interface to form a two-dimensional electron gas (2DEG). The main interest of these structures is that the carriers of the 2DEG, as they are not submitted to impurity scattering, can achieve high mobilities and velocities. The principal drawback of the conventional MODFET structure is that, because of the gate potential, carriers increase their energy, leave the two-dimensional electron gas and deeply penetrate the GaAs layer and even the substrate. Therefore this process results in a reduction of the transconductance Gm and in a high value of the output conductance Gd. A solution to this problem is the use of an Inverted Structure in which the heterojunction is located behind the 2DEG and should reduce the penetration of the carriers into the buffer layer and improve the performance of the device. A simulation was carried out to study and analyze the various , physical aspects and the potential performances of both conventional and inverted structures I-GaAs/N^+-AlGaAs.

THE MODEL

As not stationary dynamic effects(overshot, ballistic effect) occur in these submicronic structures the method used must take these effects into account. A Monte-Carlo method was used, allowing as to account for not stationary dynamic phenomena, then a two-dimensional solution of poisson equation was associated so as to account for the bidimensional spatial effects which can become important when the dimensions of the devices studied are very small. Nevertheless, this model, like other models [1] consider the two-dimensional electron gas carriers as bulk carriers and do not take into account the effects linked to the energy quantification in the conducting channel. There exist more sophisticated models [2] [3] which partially take into account the quantum effect and necessitate longer computing times but apparently do not yield very different results.

CONVENTIONAL MODFET

Conventional MODFET structure consist mainly of a GaAs buffer layer on which is grower first an undoped active GaAs layer and then an heavily doped AlGaAs layer on which the gate is deposited. The 2DEG is located in the undoped GaAs layer, at the GaAs/AlGaAs interface. We simulated a 0.3 μm gate length (Lg) conventional structure with a 400Å thick (a) AlGaAs layer, doped to 10^{18} at/cm^3 (Nd), the aluminium composition of which being 30%.
For this structure the theoretical pinch-off voltage is 1.5 Volt, given by the relation :

$$Vp = qN_d a^2/2\varepsilon + \Delta E_c$$

where ΔE_c is the conduction band discontinuity.

In figure 1, are presented the variations of the drain current Ids, the transconductance Gm and cut-off frequency Fc versus the internal gate voltage Vgs-Vbi (where Vbi is the built-in potential) at room temperature.
These curves exhibit two different features depending on gate bias. For gate voltage lower then $-qN_d a^2/2\varepsilon = -1.2$ V, the AlGaAs upper layer is completely depleted , then the drain current is only due to the 2DEG and the transconductance Gm varies almost linearly with gate voltage. For gate voltage greater than -1.2 V, current conduction occur in the AlGaAs layer givingrise to a " MESFET like " contribution. The Gm variation versus gate voltage is then reduced due to the low transport properties carriers mobility of the highly AlGaAs.

In this voltage range, the charge controle is carried out either on the 2DEG and on gate capacitance Cgs the AlGaAs N^+ carriers, thus increasing the gate capacitance Cgs and decreasing the cut-off frequency Gm/2πCgs. All these aspects leads to a cut-off frequency variation versus gate voltage showing a pronounced maximum around -1.2 Volt. For this gate bias which can be looked at an optimum polarisation point, the device exhibit a 400 mS/mm transconductance associated with 110 Ghz cut-off frequency.

At liquid nitrogen temperature, for the same polarisation point (Vgs-Vbi = -1.2 V) , the performances increase to Gm = 600 mS/mm and Fc = 150 Ghz due to better transport properties of the GaAs at low temperature.

As concerned with the output conductance Gd, for the same gate bias, it varies from 20 mS/mm at room temperature to 28 mS/mm at 77K. These relatively high

values are strongly connected with carriers injection in the GaAs buffer layer in the gate to drain region. This is one of the main inconvenience of conventional MODFET structure, an other one is related to the fact that the gate potential acts not only on the 2DEG but also on AlGaAs layer ionized impurities, thus reducing the gate efficiency and the transconductance values.
A solution to this problem is the use of inverted structure.

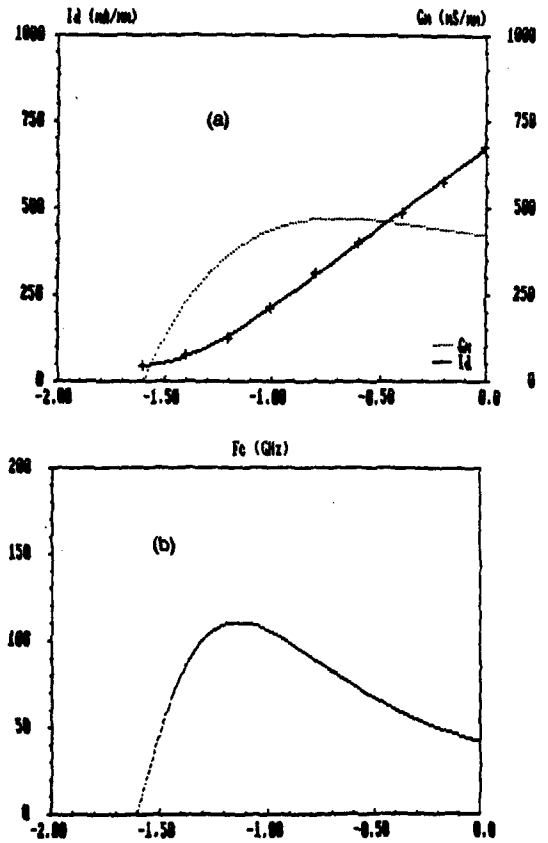


Figure 1 : Gm (a) and Fc (b) versus internal Vgs for classical MODFET

INVERTED MODFET STRUCTURE

In the inverted structure, the gate is deposited on an undoped GaAs layer which is grown on a highly doped AlGaAs layer. The source and drain contacts are taken on a top GaAs layer heavily doped in order to reduce access resistances and completely recessed so that the gate is on the undoped GaAs layer.

We simulated structures with a 1.5μm source-drain spacing and a 1.1μm long by 0.1μm deep recess and we studied the evolution of performances when modified the thickness of the GaAs undoped layer, the thickness, the doping level or aluminium composition of the $Al_xGa_{1-x}As$ layer, the gate length or the temperature.

In all case, the gate being on the undoped material, the gate potential acts directly on the channel carriers and not on the carriers and the ionized impurities as it is the case in a conventional structure. The gate action is thus much more efficient and results in much higher transconductance values than for conventional structures. Thus for a 0.3μm gate length, a 250Å AlGaAs layer, the maximal transconductance value is approximately 750 mS/mm at room temperature.

Influence of the GaAs thickness a and of temperature

If, for the same structure the GaAs layer thickness is reduced, the Gm maximal value changes to 1350 mS/mm. These results are given in figure 2 where , for each structure, the Gm evolution versus the gate voltage is represented (internal potential, i.e when taking the Schottky potential into account). It can be seen that for a = 125Å, the Gm evolution versus Vgs is very steep, which might lead to interesting developments in high speed logics.
On the other hand the latter structure, simulated at 77K gives a Gm maximal value of approximately 2200 mS/mm. These values are of the same order of magnitude as the experimental values obtained by CIRILLO [4] for an similar structure.

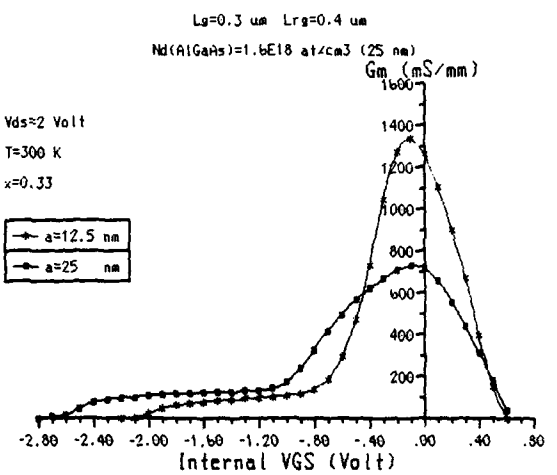


Figure 2 : Gm versus internal Vgs for 2 structures with different active layer thickness

Influence of the gate length Lg

Figure 3a shows the Gm evolution versus the gate voltage for three equivalent structures where only the gate length is modified. It can be seen that the longer the gate, the more efficient the gate control and the higher the transconductance. But the capacitance Cgs increases with Lg and therefore the cut-off frequency Fc decreases when Lg increases (Fig 3b).

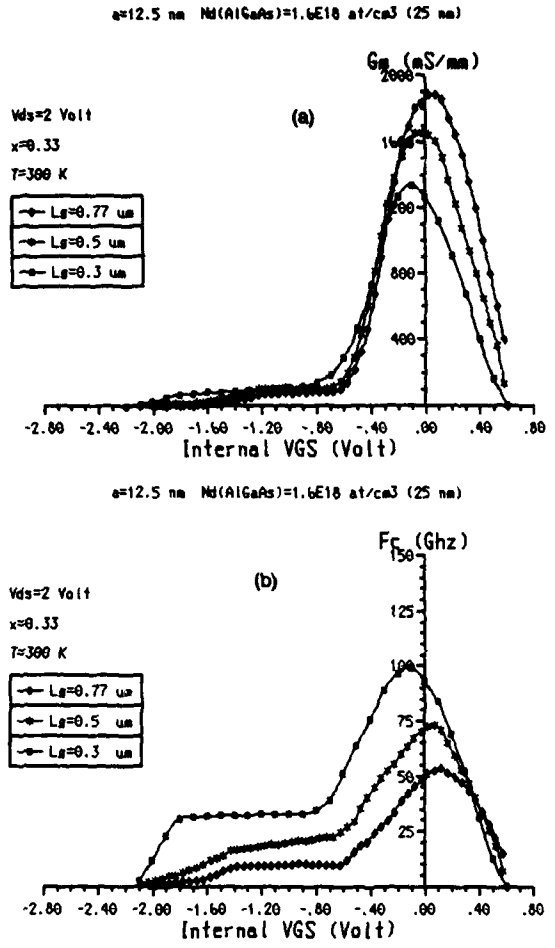


Figure 3 : Gm (a) and Fc (b) versus internal Vgs for 3 structures with different gate length

Influence of thickness and doping of the AlGaAs layer

We also simulated structures where the AlGaAs layer parameters(thickness and doping) were modified. It can be seen in figure 4a and 4b that these parameters have little influence on the static characteristics of the component. In particular, the transconductance maximal value remains the same, only the characteristic Gm(Vgs) is shightly modified and for a given Vgs voltage , Gm keeps high values when the AlGaAs layer is very thin or highly doped.

Influence of the $Al_xGa_{1-x}As$ layer composition x

We simulated structures where only the aluminium composition of the AlGaAs layer was modified, which results in a modification of the conduction band discontinuity at the GaAs/AlGaAs heterojunction. It can be seen in figure 5 that the Gm maximal value increases with the composition x insofar as x is inferior to 0.3 and that for higher composition Gm does not change. On the other hand, the output conductance strongly decreases with x and this is due to the potential barrier increase at the GaAs/AlGaAs heterojunction, thus reducing the penetration into the substrate.

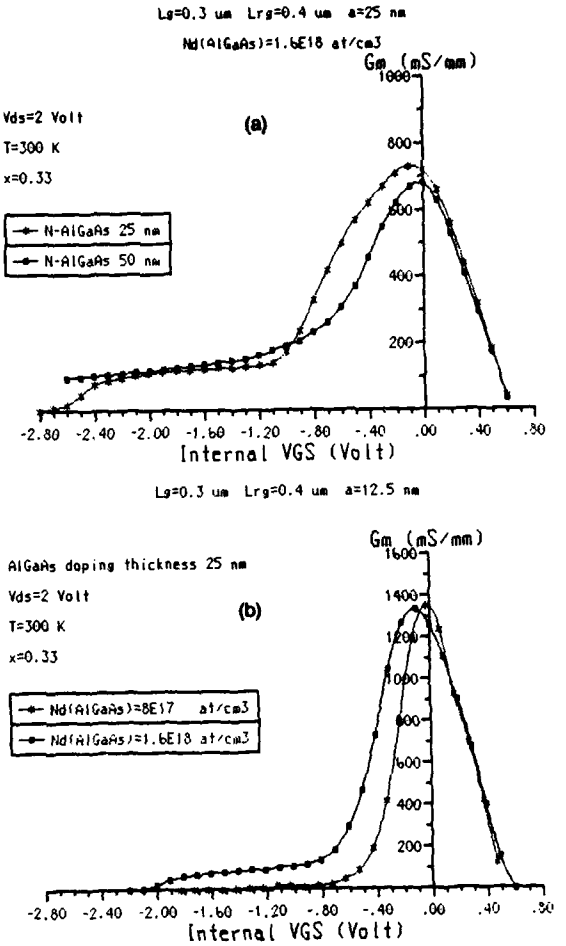
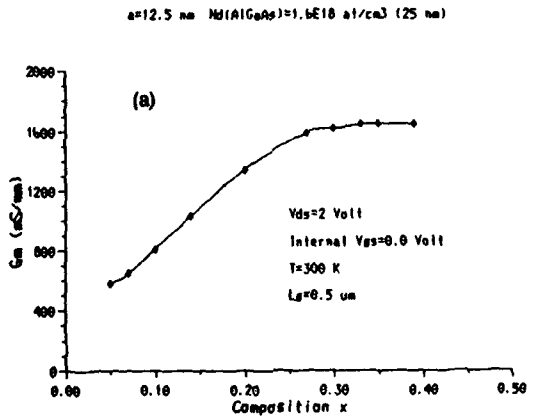


Figure 4 : Gm versus internal Vgs for 2 different AlGaAs layer thicknesses (a) and doping (b)



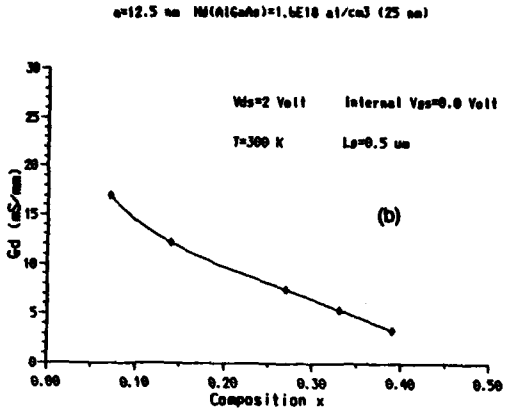


Figure 5 : Gm (a) and Gd (b) versus internal Vgs the Al composition of the AlGaAs layer

Conclusions

The MODFET inverted structure could lead to very interesting developments. In particular, the use of thin GaAs layers, associated with appropriate thickness and doping level for the AlGaAs layer allows the obtention of much higher transconductance and lower output conductance as compared to conventional MODFET structures.

Finally, the choise of the gate length must achieve a compromise between the transconductance value and the maximal cut – off frequency.

REFERENCES

[1] T. WANG ,K. HESS
calculation of the electron velocity distribution in HEMT using an ensemble Monte – Carlo method
J.Appl. Phys. 57,p 5336 (1985)

[2] M. MOTIZAWA ,A. YOSHII ,L. YOKOYAMA
Modeling for an AlGaAs/GaAs device using Monte – Carlo simulation
IEEE El. Dev. Lett., EDL – 6,7, p 332 (1985)

[3] V. RAVAIOLI ,D.K. FERRY
MODFET ensemble Monte – Carlo model including the quasi two dimensional electron gas
IEEE Trans. on Electron. Dev., ED – 33,5,p 677 (1986)

[4] N. CIRILLO
Inverted GaAs/AlGaAs MODFET with extremely high transconductance
IEEE El. Dev. Lett., EDL – 7,2, p 71 – 74 (1986)

The Particle Simulation of self-aligned GaAs MESFETs with a Submicrometer Gate-length

Yoshinori YAMADA, Shiroh IKEDA, and Naomasa SHIMOJOH

Department of Electrical Engineering and Computer Science
Kumamoto University, Kumamoto 860, JAPAN

The self-aligned GaAs MESFET with 0.25 μm gate-length has been analyzed by the two-types of particles simulator(TPS) and the regional simulator(RES) which are useful for drastic reduction of the number of particles. The numbers of particles employed in the TPS and RES are less than that in the conventional simulator by factors of about 8 and 18 - 54, respectively.

1. INTRODUCTION

It is said that a multi-particle simulator is a good tool to investigate nonstatic transport in GaAs MESFETs with a short channel. Awano et al. applied it to an entire region of a self-aligned MESFET[1]. Their simulator is called a full simulator(FUS) hereafter.

The self-aligned MESFET consists of highly-doped n^+-regions and a lightly-doped n-region. In the FUS a particle denotes an ensemble of electrons and an electric charge in a particle is same for all the particles. Thus, the n^+-regions required a large number of particles in comparison with the n-region. For example, the number of particles employed by Awano et al. reached about 65 000. Among the particles only about 3 000 particles in the n-region played an essential role for description of the nonstatic effect. The others were employed in order to describe the near Ohmic transport in the n^+-regions.

As the FUS requires a large number of particles and large amounts of CPU memory and CPU time, it is not considered to be the best tool for designs. Some ideas which are useful for reduction of the CPU memory and CPU time have been presented. The one is to use ballance equations[2 - 5]. The another is a regional simulator[6 - 8].

The purpose of the present work is to describe a two-types of particles simulator(TPS) and our regional simulator(RES). Both of them are useful for reduction of the number of particles. The TPS is based on a different idea from the previous ones.

2. TWO-TYPES OF PARTICLES SIMULATOR(TPS)

Fig. 1(a) shows a schematic drawing of the self-aligned structure used in the present simulations. A substrate is not considered.

Fig. 1(b) explains a concept of the TPS. It employs two different types of particles and makes a charge in a particle in the n^+-regions larger than that in the n-region by a factor of η (>1). When the charges in a particle in the n- and n^+-regions are denoted by q and q^+, respectively, q^+ is equal to ηq. The FUS corresponds to a case of $\eta = 1$. A choice of the value of η depends on how much degree one desires to reduce the number of particles and permits inaccuracy of the calculated results. The TPS applies the particle simulation to an entire of the device structure shown in Fig. 1(a). In general, an increase of the charge in a particle causes instabilities or fluctuations in the numerical calculation. It will be shown, however, in the later section that the TPS does not cause large fluctuations.

2.1. Number of Particles

The number of particles in the n^+-regions used in the TPS is less than that used in the FUS by a factor of η. For example, if $\eta = N_D^+ / N_D$ and $L_{SG} = L = L_{GD}$, the total number of particles used in the TPS is less than that in the FUS by a factor of $(2\eta+1)/3$. If $\eta = 10$, the factor is equal to 7. Here N_D^+ and N_D are the donor concentrations of the n^+- and n-regions, respectively.

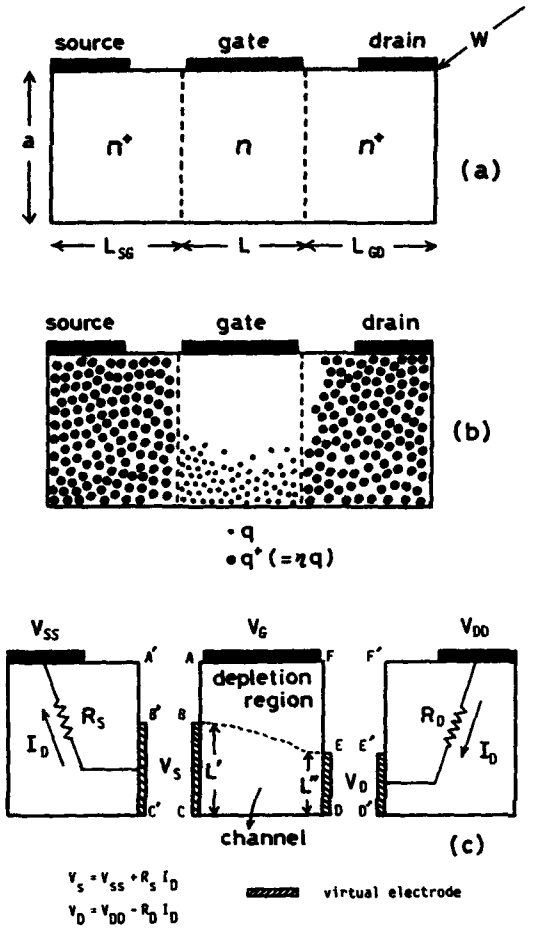


FIGURE 1

(a) A schematic drawing of a self-aligned MESFET. (b) A concept of the TPS. (c) A principle of the RES.

2.2. Algorithm Assuring Current Continuity

As the TPS uses the different types of particles in the n^+- and n-regions, it is important to consider an algorithm assuring current continuity at the junction boundaries.

When a particle with q^+ injects into the n-region, it must be replaced by η particles with q. When particles with q inject into the n^+-region, the number of the injecting particles is counted. When the number reaches to η, they are replaced by a particle with q^+. In addition to the exchanges of the particles their positions and wave vectors must be also determined from the current continuity. They are successfully determined from the position and wave vector distribution-functions of the injecting particles by using random numbers. The detailed description will be published elsewhere.

3. REGIONAL SIMULATOR(RES)

Fig. 1(c) explains a principle of our RES. The RES analyzes the n-region by the multi-particle simulation and replaces the n^+-regions by series resistances R_S and R_D. The "virtual electrodes" shown by a shaded portion in Fig. 1 (c) are set on the junction boundaries. During the simulation the density of particles in the virtual electrodes is kept constant as well as in the FUS.

3.1. Number of Particles

Naturally the number of particles required in the RES is less than that in the FUS by a factor of about $(2\eta+1)$, if $\eta = N_D^+ / N_D$ and $L_{SG} = L = L_{GD}$. The factor is equal to 21 if $\eta = 10$.

3.2. Virtual Electrodes

Strickly speaking, the particles in the virtual electrodes should not be at thermal equilibrium. However, as N_D^+ is fairly high, we have assumed that the particles are at thermal equilibrium.

The widths L′ and L″ are considered to be equal to the channel widths near the junction boundaries. Consequently they depend on a geometry of the depletion region.

3.3. Carrier Concentration

The density of particles within the virtual electrodes corresponds to the donor concentration and is kept constant. Thus, it does not directly correspond to the carrier concentration. Consequently the carrier concentration has been approximately estimated from the density of particles multiplied by a correction factor. The correction factor, which is a function of the channel distance, is defined as a ratio of the real carrier concentration to the density of particles at thermal equilibrium. The real carrier concentration along the channel at thermal equilibrium may be easily estimated from the one-dimensional Poisson's equation. If the virtual

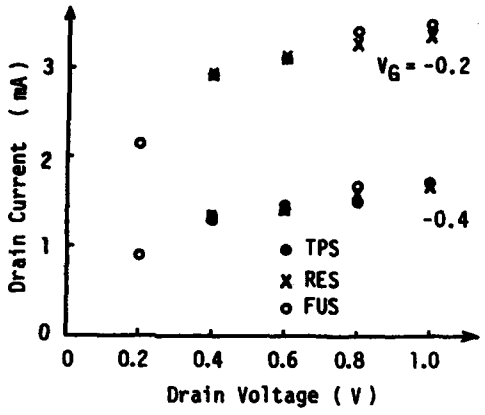


FIGURE 2
Calculated I_D - V_D characteristics

electrodes are set a little bit inside of the n^+-regions, the correction factor may not be needed.

4. NUMERICAL RESULTS

The device parameters used in the simulations are as follows.: L=0.25 μm, a=0.1 μm, N_D=7x10^{16} cm^{-3}, and N_D^+=10N_D. The TPS has been carried out by using 7 411 particles and the RES by using 1200 or 3600 particles. The numbers of particles in the TPS and RES are less than that in the FUS by factors of 8 and 18 - 54, respectively.

Fig. 2 shows the calculated I_D - V_D characteristics, where I_D is the drain current and V_D the drain voltage. The V_G is the gate voltage which includes a barrier height. It is seen that the

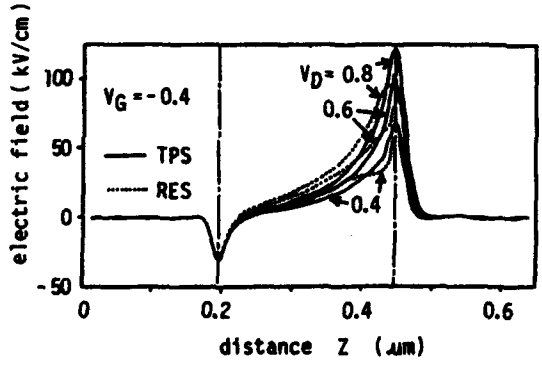


FIGURE 3
Average electric field distribution

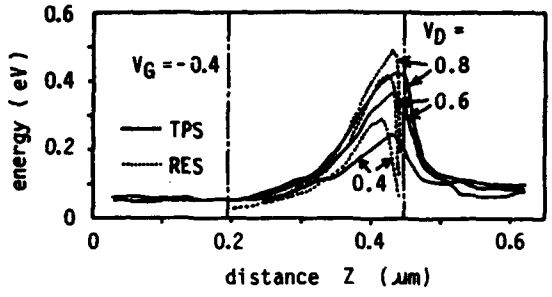


FIGURE 4
Average energy distribution

calculated values by the TPS and RES excellently agree with those by the FUS.

Now let us compare the electric field, energy and velocity distributions in the device obtained by the TPS and RES. In the following figures, the calculated results by the TPS and RES are denoted by solid and dotted lines, respectively.

Fig. 3 shows the average electric field distributions parallel to the channel at the bottom of the channel. Owing to the use of the correction factor, the height of the reflecting barrier at the source junction by the RES fairly agrees with that by the TPS. If the factor is not used, the barrier may be higher. The figure tells us that the electric field by the RES is strong at the middle of the channel and is weak near the drain virtual elctrode, in comparison with that by the TPS. It is the reason that the RES does not consider a high resistive and narrow region of the drain n^+-region.

Fig. 4 shows the average energy distributions at the bottom of the channel. Due to the assumption of thermal equilibrium in the virtual electrodes, the energy by the RES is smaller near the virtual electrodes than that by the TPS. The peak energies by the RES are higher than those by the TPS because the electric field by the RES is larger at the middle of the channel than that by the TPS.

The average velocity distributions parallel to the channel at the bottom of the channel are shown in Fig. 5. The distributions for V_D=0.6 and 0.8 volts are shifted for clear illustration.

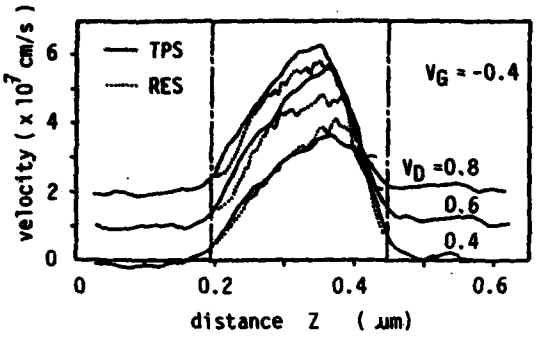


FIGURE 5
Average velocity Distribution

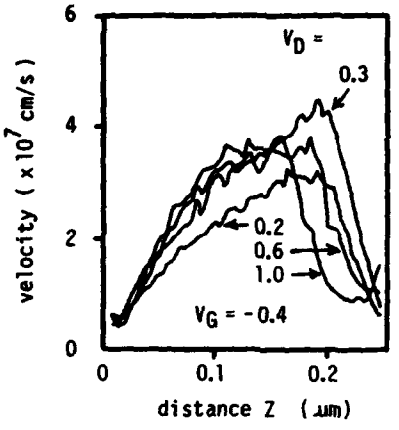


FIGURE 6

The near ballistic transport in the device is clearly observed in the n-region. As shown in Fig. 4, the energy increases with V_D and the slowing down of the velocity due to the nonequivalent intervalley scattering extends with V_D. Thus, the peak position of the velocity distribution moves toward the source n^+-region.

Fig. 6 shows a dependence of the velocity distribution upon V_D which is calculated by the RES. In addition to the shift of the peak position of the velocity distribution due to the nonequivalent intervalley scattering, as described previously, we observe that the peak velocity is maximum at V_D=0.3 volt. On the other hand,the TPS tells us that the peak velocity at V_D=0.6 volt is larger than those at V_D=0.4 and 0.8 volts. Due to our poor computer resources, enough data could not be obtained by the TPS. According to our experience, it is not easy to achieve good statistical convergence of the velocity distribution in comparison with the other distributions. Thus, we cannot conclude in this work that the peak velocity decreases in the saturation region of the I_D - V_D characteristics as V_D increases.

5. CONCLUSIONS

It has been found that both the TPS and RES give the reasonable I_D - V_D characteristics. The numbers of particles employed in the TPS and RES are less than that in the FUS by factors of 8 and 18 - 54, respectively. The influences of the assumptions used in the RES on the electric field, energy and velocity distributions have been elucidated. The dependence of the peak velocity on the drain voltage has been discussed.

REFERENCES

[1] Awano, Y., Tomizawa, K., and Hashizume, N., IEEE Trans. Electron Devices, Vol. ED-31, No. 4(1984)448

[2] Blotekjar, K., IEEE Trans. Electron Devices, Vol. ED-17, No. 1(1970)38

[3] Cook, R. and Frey, J., IEEE Trans. Electron Devices, Vol. ED-29, No. 6(1982)970

[4] Curtice, W. R. and Yun, Y., IEEE Trans. Electron Devices, Vol. ED-26, No. 8(1981) 954

[5] Carnez, B., Cappy, A., Kszynski, A., Constant, E., and Salmer, G., J. Appl. Phys., Vol. 51, No. 1(1980)784

[6] Furuta, T., Tomizawa, M., Yokoyama, K., and Yoshii, A., *IECE of Japan, Techn. Report*, Vol. CAS-86, No. 149(1986)87

[7] Park, Y.-J., Navon, D. H., and Tang, T.-W., IEEE Trans. Electron Devices, Vol. ED-31, No. 12(1984)1724

[8] Nguyen, P. T., Navon, D. H., and Tang, T.-W., IEEE Trans. Electron Devices, Vol. ED-32, No. 4(1985)783

Session C1.2

Compound Semiconductors Technology II

Chairman: A. Cetronio

Monday, September 14, 1987

REDISTRIBUTION OF ION-IMPLANTED MERCURY DURING RAPID THERMAL ANNEALING OF $Ga_{0.47}$ $In_{0.53}As$ and InP

J.H. Wilkie and B.J. Sealy

Department of Electronic and Electrical Engineering
University of Surrey
Guildford, Surrey GU2 5XH, England

Rutherford backscattering (RBS) has been used to study the redistribution of ion-implanted mercury and the reordering of implantation damage during Rapid Thermal Annealing (RTA) of $Ga_{0.47}$ $In_{0.53}As$ epilayers and bulk InP. Implantation at 200°C prevents amorphous layer formation and reduces the extent of Hg redistribution in both materials. Si_3N_4 and phosphosilicate glass (PSG) are compared as encapsulants for InP; it is suggested that silicon indiffusion enhanced by radiation damage may be the reason why p-type activity has only been seen for 200°C implanted material.

1. INTRODUCTION

Ion implantation of both donors and acceptors is an attractive method of forming shallow and abrupt p-n junctions in the III-V semiconducting materials GaInAs and InP. Implantation and annealing studies of the acceptors Be, Mg, Zn and Cd [1-4] have shown that these species often redistribute markedly during post-implant annealing, leading to non uniform junction depths and erratic electrical behaviour.

Room temperature Hg^+ implantation has recently yielded p-type activity in both GaInAs and InP following long time furnace annealing [4,5] and 200°C implants have been activated in InP using RTA [6]. The redistribution of implanted Hg during post-implant annealing has been studied by Secondary Ion Mass Spectrometry (SIMS) for the case of InP [5], but is not accessible to this analysis technique in GaInAs because of 'matrix interference effects' [4]. An alternative method of surface analysis is RBS which is sensitive to the presence of Hg in both of these materials and gives simultaneous information about the nature and extent of crystalline damage.

2. EXPERIMENTAL

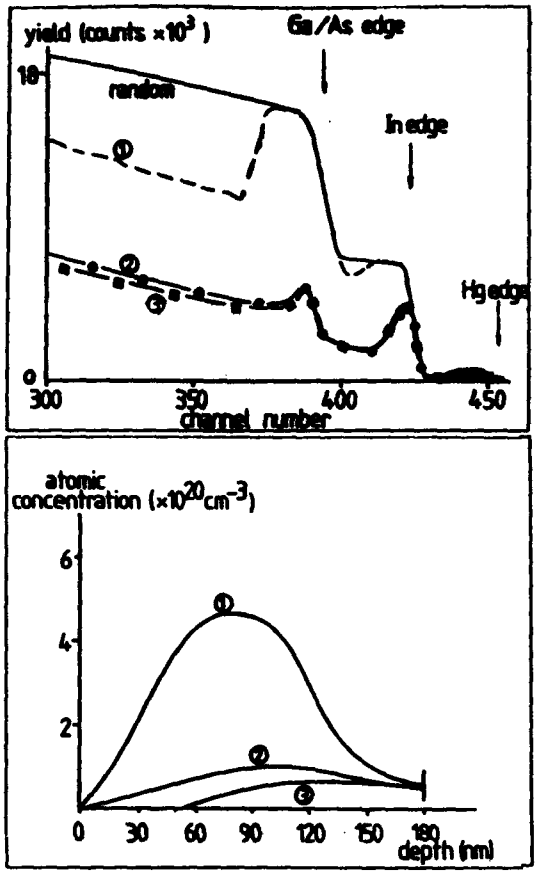
In the present work, 100 keV Hg^+ implants of 1-5 x 10^{15} cm^{-2} were made into bare InP substrates and 'lattice matched' LPE grown GaInAs/InP epilayers held either at room temperature or 200°C. Following implantation, samples were subjected to RTA at temperatures between 500° and 800°C for times of 30, 60 or 120 seconds using a variety of surface protection techniques. RBS (1.5 MeV, He^+) was then used to follow the redistribution of Hg and the removal of implantation damage that had occurred during the encapsulation/annealing cycle. This data was correlated to electrical measurements.

3. RESULTS

3.1. GaInAs

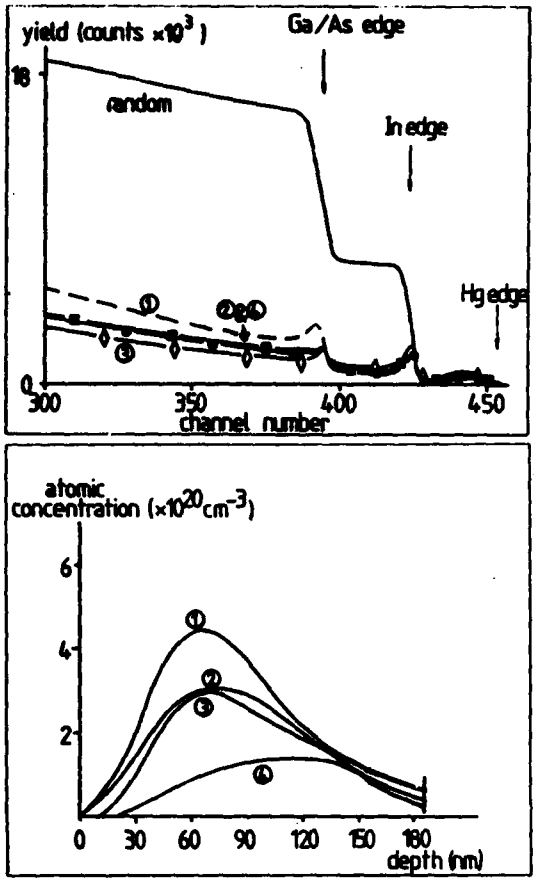
RBS results for Hg^+ implanted into GaInAs are shown in Figures 1 and 2.

It may be seen from Fig. 1(b) and Fig. 2(b) that the redistribution of implanted Hg in GaInAs during RTA under a GaAs 'proximity cap' is quite different for the two implant conditions, giving peak detected Hg levels following a 700°C, 30 second anneal of ~1 x 10^{20} cm^{-3} for the RT implant and



FIGURES 1(a) and (b)

RBS spectra and derived Hg concentration profiles for 10^{15} cm^{-2}, 100 keV RT Hg implants into LPE GaInAs. (1) as implanted, (2) 600°C, (3) 700°C. All anneals 30 sec., GaAs proximity cap.



FIGURES 2(a) and (b)

RBS spectra and derived Hg concentration profiles for 10^{15} cm^{-2}, 100 keV 200°C Hg implants into LPE GaInAs. (1) as implanted, (2) 600°C, (3) 700°C, (4) 800°C. All anneals 30 sec., GaAs proximity cap.

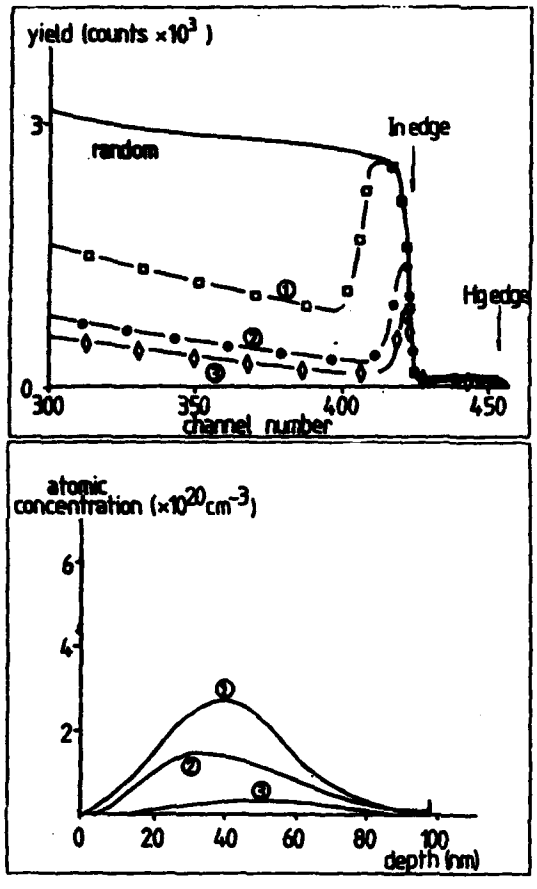
~3 x 10^{20} cm^{-3} for the 200°C implant. RTA at 800°C for 30 seconds leads to Hg being detectable only in the 200°C implant case.

3.2. InP

1 x 10^{15} cm^{-2}, room temperature Hg^+ implants into InP form an amorphous layer (Fig. 3(a)) RTA at 500° and 600°C for 120 seconds under a phosphorous overpressure provided by a 10% Sn/90% InP powder [7] regrows the amorphous layer, leaving thin damage regions near the surface. 200°C implantation of 5 x 10^{15} cm^{-2} Hg into InP suppresses amorphous layer formation although a small damage peak may be observed.

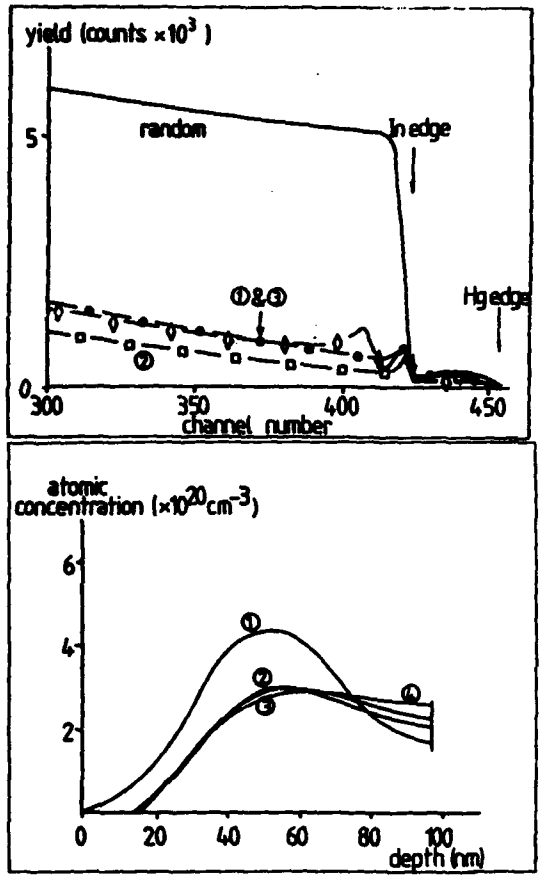
As for the RT GaInAs implants, Hg^+ implanted at RT into InP redistributes significantly during RTA at 500° and 600°C (Fig. 3(b)). Further anneals (not shown) at 700°C and 800°C, for 90 and 60 seconds respectively, led to no Hg being detectable at the 1 x 10^{20} cm^{-3} level.

Figs. 4(b) and 5(b) compare the Hg redistribution for the 200°C implant condition during RTA under two different encapsulants; 1000 Å of pyrolytic Si_3N_4, deposited at 595°C and CVD PSG deposited at 300°C. In contrast to the results for 'overpressure' annealing of



FIGURES 3(a) and (b)

RBS spectra and derived Hg concentration profiles for 10^{15} cm^{-2}, 100 keV RT Hg implants into InP. (1) as implanted, (2) 500°C, (3) 600°C. All anneals 120 sec., overpressure technique.



FIGURES 4(a) and (b)

RBS spectra and derived Hg concentration profiles for 5 x 10^{15} cm^{-2}, 100 keV 200°C Hg implants into InP. (1) as implanted, (2) as capped, (3) 700°C, 60 sec. Si_3N_4 capped.

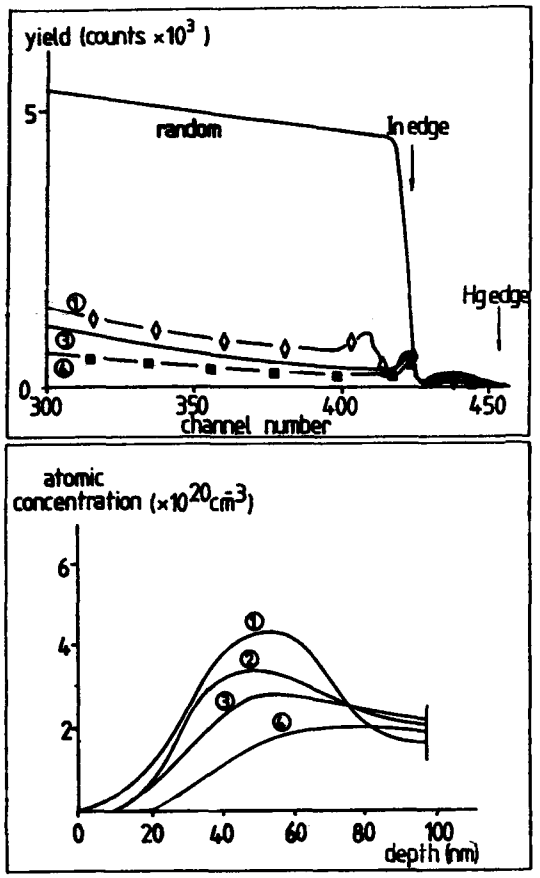
RT implants, significant levels of Hg remain detectable even after 800°C anneals, with Si_3N_4 apparently retaining slightly more Hg at the higher annealing temperature.

3.3. Electrical Results

Hall effect measurement using the Van der Pauw sample geometry has shown that both RT and 200°C Hg^+ implanted GaInAs are n-type for all annealing conditions. Similar measurement of Hg implanted InP has yielded p-type activity for 200°C implants [6] and n-type activity for room temperature implants.

4. CONCLUSIONS

Room temperature implantation of doses >10^{15} ions cm^{-2} of 100 keV Hg^+ into InP and GaInAs forms amorphous layers, regrowth of which cause significant redistribution of the implanted Hg. Implantation at 200°C suppresses this amorphous layer formation and leads to increased levels of Hg being detected in samples annealed up to 800°C. The most likely cause of n-type electrical activity seen following RTA of InP samples implanted at room temperature is considered to be



FIGURES 5(a) and (b)

RBS spectra and derived Hg concentration profiles for 5 x 10^{15} cm^{-2}, 100 keV 200°C Hg implants into InP. (1) as implanted, (2) 600°C, (3) 700°C, (4) 800°C. All anneals 60 sec. PSG capped.

implantation damage enhanced indiffusion of silicon from the Si_3N_4 and PSG used as encapsulants. N-type activity of both RT and 200°C Hg implanted GaInAs cannot be explained at present.

ACKNOWLEDGEMENTS

The authors wish to thank the SERC for funding, British Telecom Research Laboratories for the InP, Sheffield University, Department of Electronic Engineering for the GaInAs and the staff of the D.R. Chick Laboratory at the University of Surrey for assistance with the Hg^+ implantations.

REFERENCES

[1] Tell, B., Leheney, R., Liao, A., Bridges, T., Burkhardt, E., Chang, T., and Beebe. D., Appl. Phys. Lett. 44, 43 (1984).

[2] Vescan, L., Selders, J., Maier, M., Krautle, H. and Beneking. H., J. Cryst. Growth 67, 353 (1984).

[3] Wilkie, J.H., Spiller, G.T., Henning, I.D. and B.J. Sealy, submitted to J. Cryst. Growth (1987).

[4] Favennec, P., L'Haridon, H., Gauneau, M., Salvi, M., Roquais, J. and Razeghi, M. Inst Phys. Conf. Ser. No. 79 Chapter 6, presented at Int. Symp. GaAs and related compounds, Karuizawa, Japan, (1985).

[5] Favennec, P., L'Haridon, H., Roquais, J., Salvi, M., LeCleach, X. and Gouskov., L. Appl. Phys. Lett. 48, 154 (1986).

[6] Wilkie, J.H. and Sealy, B.J. Elec. Lett. 22, 1308 (1976).

[7] Wilkie, J.H. and Sealy, B.J., paper presented at ESSDERC Conf. 8-11 September 1986, Cambridge, U.K.

RAPID THERMAL ANNEALING OF Be IMPLANTS INTO UNDOPED InP

W. Häussler

Siemens AG, Research Laboratories, ZFE FKE 1
Otto-Hahn-Ring 6, 8000 München 83, F. R. G.

The anneal behavior of Be implantations into undoped InP was studied using a capless, rapid thermal annealing (RTA) process. It is shown that shallow p-type layers in InP can be obtained reproducibly making this process suitable for device fabrication. Maximum hole concentrations and minimum in-diffusion are observed for anneal durations of about 1 minute. Out-diffusion of implanted Be is identified as the main reason for low hole concentrations and can be decreased by co-implanting matrix atoms.

INTRODUCTION

Ion implantation into InP is an important technology for future optoelectronic integrated devices. Be [1][2], Mg [3][4], Zn [3] and Cd [4] have been used for acceptor implantation. It has been established, however, that these acceptors exhibit strong in-diffusion during furnace-annealing. In-diffusion should be reduced by rapid annealing. Rapid annealing has been applied to low-dose Be implants into InP:Fe [5], indicating absence of in-diffusion, to Mg implants into InP:Fe [6], showing strong in-diffusion, and to Zn-implants [7] with little in-diffusion for undoped InP and strong in-diffusion for Fe-doped InP.

In this paper results of a systematic study on rapidly annealed Be implantations into undoped InP are reported. The implanted layers were studied under various anneal conditions. Substrate effects were also investigated by comparing implants into LEC-, LPE- and MOCVD-grown InP. Furthermore, preliminary results of additional implantations of P- and In-ions are reported.

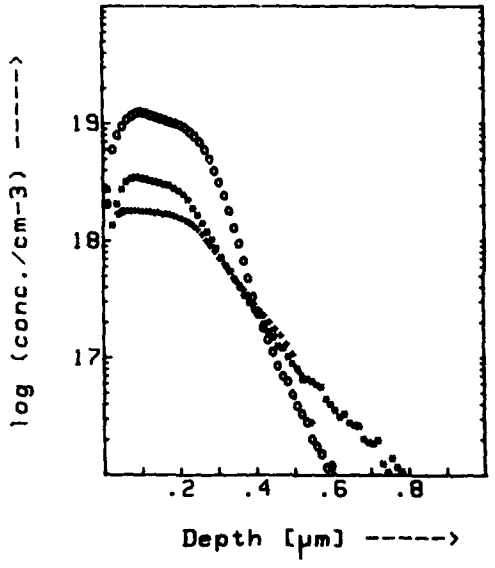
EXPERIMENTAL

Be was implanted with energies between 20 and 100 keV and doses ranging from 5.6E+13 to 2.0E+15 cm^{-2}. A graphite strip heater was used to achieve rapid heating of the sample of up to 100 °C/sec. Decomposition of the InP surface was avoided in a PH_3/H_2 atmosphere. No dielectric cap was employed and the InP surface was mirror-like after annealing. Anneal temperatures ranged between 750 and 850 °C with hold-times between 6 seconds and 4 minutes. Implanted InP layers were non-intentionally doped with electron concentrations below 1E+16 cm^{-3}.

The distribution of Be atoms after the anneal was measured using SIMS. Multiple measurements on each sample gave a dose accuracy of about 5 %, so that precise out-diffusion data could be obtained.
The distribution of holes after annealing was measured with an electrochemical C(U)-profiler [8]. Measurement values were corrected to include the effect of series resistance.

RESULTS

A typical annealing result for a dual energy Be implant of 20 and 50 keV into LEC-InP is shown in Fig. 1. The two SIMS curves measured before and after the anneal show that substantial amounts of Be have been lost. The loss amounts to 71 % for a dose of 2.75E+14 cm^{-2} and an anneal at 850 °C for 1 minute. Even for a 6" anneal at 800 °C out-diffusion is 66 % . Out-diffusion depends on implantation dose, but is still severe for smaller doses (Table 1).



SIMS profiles for a 20+50 keV Be-implant before (O) and after (•) an anneal at 850 °C-1' showing substantial out-diffusion of the implant. Total implanted dose was 2.75E+14 cm^{-2}. The resultant hole concentration profile is also indicated (+). Electrical activation is better than 50 % .

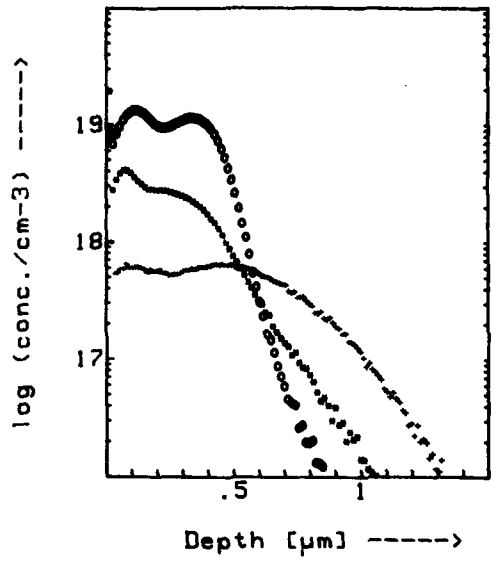
TABLE 1: Be out-diffusion as a function of dose

implanted dose	out-diffusion
5.60E+13 cm^{-2}	57 %
1.35E+14 cm^{-2}	71 %
2.75E+14 cm^{-2}	71 %
2.00E+15 cm^{-2}	94 %

Apart from out-diffusion, there is only minor in-diffusion. For longer annealing periods in-diffusion increases.

The hole concentration profile shown in Fig. 1 indicates that a minimum of about 50 % of the Be atoms is electrically active resulting in a maximum hole concentration of about 2E+18 cm^{-3}. Electrical activation approaches 100 % for lower Be concentrations.
Electrical activation is very low for a 6" anneal at 850 °C. Results of 30" anneals are comparable to 1' anneals. Annealing periods of about 1 minute at 850 °C therefore represent optimum conditions.

Hole concentration profiles of annealed Be-implants into LEC-InP have been very reproducible for several annealing and implantation runs. However, this is not necessarily so for implantations into LPE-InP, where even with RTA significant in-diffusion of Be is observed occasionally (Fig. 2). Surprisingly, by reducing the anneal temperature to 800 °C in-diffusion can be reduced and the maximum hole concentration can be increased. No similar effect is seen with LEC-InP.
Results for Be implants into MOCVD-grown InP were comparable to implants into LEC-InP.
These data indicate that annealing parameters may have to be modified for differently grown InP to obtain optimum results.



SIMS profiles for a 30+100 keV Be-implant before (O) and after an anneal at 850 °C showing that the implanted material can have a significant influence on Be-redistribution : strong in-diffusion is observed in this case for LPE-InP (+); very little in-diffusion is observed for LEC-InP (*). Implanted dose has been 5.0E+14/cm².

Effect of a 400 keV In implantation on a Be-implanted sample. Implantation and annealing parameters and plotting symbols are identical with Fig. 1.

In an attempt to increase the maximum hole concentration additional P and In implantations were performed. Both ions were implanted to give a maximum concentration of 1.0E+19 cm⁻³.
Fig. 3 shows the annealing result for an additional 400 keV In-implant performed after a 20+50 keV Be-implant. By comparing the Be-atom profile after annealing with the corresponding profile in Fig. 1, it becomes evident that Be-out-diffusion is reduced and amounts to 45 % .

In addition, the Be profile is significantly broader in the upper concentration range. Both effects almost double the hole dose.

Similar profiles are obtained for an additional 260 keV P-implant, which (presumably) matches the 50 keV Be-implant: again the hole dose is doubled. The enhancement in hole concentration is significantly larger than has been observed for a Zn+P implant [7].
These results clearly show that higher hole concentrations can be obtained, if out-diffusion is reduced. However, for a 1 minute anneal at 850 °C, this is accompanied by a broadening of the doping profile, which might be attributed to the formation of an amorphous layer by the heavy-ion implant [9]. A further investigation of this effect is underway.

CONCLUSIONS

Shallow p-type layers in undoped InP can be obtained by Be implantation in conjunction with rapid thermal annealing.

The most prominent effect occuring during annealing is Be-outdiffusion with almost no in-diffusion. The reason for this is not known at present.

Co-implanting matrix atoms reduces out-diffusion, but this may simply be due to amorphous layer formation. A high-dose co-implantation of Ar, however, does not produce a similar effect.

More research is necessary, if an understanding beyond that required for device fabrication is desired.

ACKNOWLEDGEMENTS

The author would like to thank his colleagues Dr. Treichler for his expert SIMS measurements and J. Müller for his assistance with the electrochemical profiling measurements, as well as H. Kranz of the Fraunhofer-Institute, Munich for Be implantations and friendly discussions.

REFERENCES

[1] J. P. Donnelly, Nucl. Instr. & Methods, 182 (1981), p. 553

[2] M. Gauneau et al., J. Appl. Phys., 57 (1985), p. 1029

[3] T. Inada et al., J. Appl. Phys., 52 (1981), p. 6623

[4] U. König et al., J. El. Materials, 14 (1985), p. 311

[5] A.N.M.M. Choudhury et al., Appl. Phys. Lett. 43 (1983), p. 381

[6] H. Kräutle, Proc. ICSICT 1986, p. 214

[7] M. Djamei et al., paper presented at ESSDERC 1985

[8] T. Ambridge and M.M. Factor, Inst. Phys. Conf. Ser. No. 24 (1975), p. 320

[9] J.H. Wilkie and B.J. Sealy, paper presented at ESSDERC 1986

ACCUMULATION OF IMPLANTED HYDROGEN AT THE SUPERLATTICE/SUBSTRATE INTERFACE

J. M. ZAVADA

USARDSG (UK)
London NW1 5TH
UK

R. G. WILSON

Hughes Research Laboratories
Malibu, CA 90265
USA

S. W. NOVAK

C. Evans & Associates
Redwood City, CA 94063
USA *

A GaAs-AlAs superlattice grown on an undoped GaAs substrate has been implanted with 300 keV protons to a fluence of 1E16/cm2. The implanted hydrogen and the deposited Al atoms have been depth profiled using secondary ion mass spectrometry (SIMS). Measurements of the as- implanted sample show that the hydrogen depth distribution is similar to that found previously with proton implants into bulk GaAs. Near the surface, the Al depth profile exhibits an alternating pattern characteristic of the superlattice compositional pattern. Annealing the sample, at temperatures from 300 to 700 C for a period of 20 minutes, causes the hydrogen to redistribute itself both towards the surface and deeper into the crystal. However, rather than diffusing into the substrate as in bulk GaAs, the SIMS profiling shows the hydrogen stopping and accumulating at the superlattice/substrate interface. The hydrogen density at the interface is quite significant and, even after a 700 C anneal, is still above 1E18/cc.

1. INTRODUCTION

The synthesis of high quality III-V semiconductor superlattices has intitiated the development of a new class of materials with important consequences in the area of optoelectronics. These multilayer structures have been used in a variety of applications including optical, electronic, and microwave devices. Many of these applications require further processing, especially with the use of ion implantation, to produce necessary electrical insulation, optical index definition, or layer mixing. Proton implantation of GaAs multilayered structures has been a valuable processing technique for such purposes and has yielded optical waveguides [1], coupled laser arrays [2], millimeter wave mixers [3], and photodiodes [4].

Research in the past few years has shown that protons are very active entities in semiconductors and are capable of producing optical [5] and electronic effects [6]. Since these effects may interfere with device performance, it is important to understand the behavior of protons in potential device structures. In this paper we report on the distribution of protons implanted into a GaAs-AlAs superlattice. The depth distributions of both the implanted hydrogen and the deposited Al atoms are determined using secondary ion mass spectrometry (SIMS). The changes in the hydrogen distribution due to furnace annealing are also investigated and compared to the behavior of implanted hydrogen in bulk GaAs.

* Portions of this work were supported by the US Army Research Office.

2. EXPERIMENTAL CONDITIONS

The superlattice used in this study consisted of alternating layers of GaAs and AlAs grown on an undoped, semi-insulating GaAs substrate. The total thickness of the superlattice was estimated to be 6 micrometers with the individual layers being 73 A for GaAs and 40 A for AlAs. From these thicknesses, the average aluminum concentration of the superlattice was 0.354. Growth conditions relating to the superlattice and characterization of the layer parameters have been previously reported [7].

The superlattice was implanted at room temperature with 300 keV protons to a fluence of 1E16/cc. The ion beam was incident at an angle of approximately 8 degrees from the 100 direction. After implantation the sample was cleaved into sections which were annealed in a flowing gas furnace for 20 minutes. The annealing temperatures ranged from 300 C to 700 C. The annealed samples and an as-implanted piece were then examined using SIMS. Both the implanted hydrogen atoms and the deposited Al were depth profiled. The SIMS measurements were made using Ce ion sputtering and the background-subtracted sensitivity for protons in GaAs was approximately 1E17/cc. The depth accuracy achieved in these measurements was about 7 %.

3. EXPERIMENTAL RESULTS

Figure 1 shows six SIMS profiles that were obtained in this study. On this depth scale the Al concentration appears nearly constant to a depth of about 6 micrometers after which it drops sharply to zero. This fall-off is taken as the location of the superlattice/substrate interface and agrees with the estimated thickness of the superlattice. Near- surface SIMS profiling with a much finer depth resolution did reveal an oscillatory pattern to the Al concentration with a period consistent with that measured for the layers using x-ray diffraction.

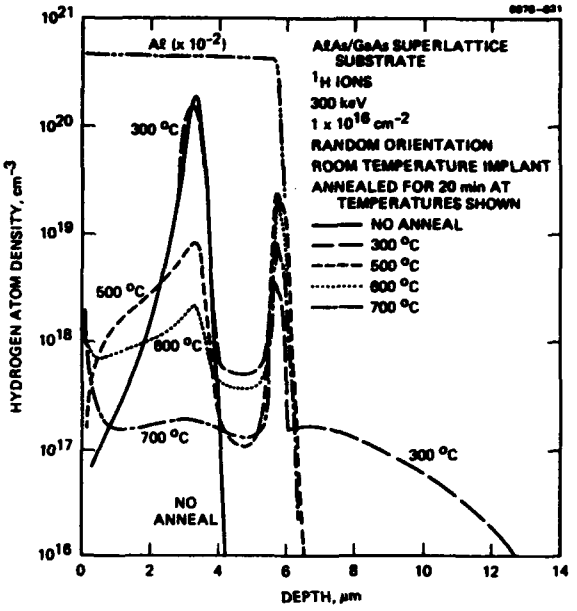


FIGURE 1

The as-implanted hydrogen depth profile is similar to that obtained for hydrogen implanted into bulk GaAs [8]. However, the peak of the distribution is deeper than that found for bulk GaAs. The SIMS profile for the sample annealed at 300 C shows that the hydrogen has started to redistribute itself. There is a slight movement towards the surface but most of the redistribution is deeper into the crystal. This is very similar to the behavior of hydrogen in bulk GaAs. However, with the implantation into a superlattice, a significant fraction of the hydrogen atoms stop and accumulate at the superlattice/substrate interface. There is a tail to the distribution that extends into the substrate but the majority of migrating hydrogen atoms are collected at the interface. The hydrogen density in this region is greater than 10% of that at the peak of the distribution.

The SIMS profile for the sample annealed at 500 C continues this development. The upward movement of the hydrogen is more pronounced and has nearly reached the surface. The density of hydrogen at original projected range has

decreased to less than 10% of its orignal value. These changes are again similar to those occurring in hydrogen implants into bulk GaAs. The density of hydrogen at the interface has increased to a level of more than 1E19/cc and now exceeds the density of the hydrogen at the original peak of the distribution.

With annealing at 600 C some of the hydrogen has reached the surface and escaped. The concentration at the projected range is about 1% of the original amount which again agrees with the behavior in the bulk GaAs. The hydrogen density at the interface is still at a level of 1E19 /cc and shows very little change from the previous anneal. At this stage most of the residual hydrogen is in the interfacial region.

Further annealing at 700 C continues this trend with most of the hydrogen concentrated at the interface. There is also a near-surface component of the distribution that has not yet escaped from the crystal. The density at the original peak is now greatly reduced, about 1/10% of the as-implanted density. At the interface the hydrogen density shows some reduction but is still greater than 1E18/cc.

4. CONCLUSIONS

The results of these SIMS measurements show that the distribution and migration of hydrogen implanted into a GaAs-AlAs is very similar to that found with implants into bulk GaAs. The projected range and general movement of the hydrogen with furnace annealing is nearly the same. However, the presence of the superlattice/ substrate interface has a major effect on the migration of the hydrogen deeper into the crystal. The hydrogen atoms are impeded by the interface, are collected there, and remain in that region to high density levels, even after annealing at 700 C. This accumulation process may be the result of lattice strain at the superlattice/substrate boundary or may be due to dislocations and imperfections arising during the initial stages of epitaxy. In an optical study of a related superlattice [9], a graded optical index in the interfacial region was required in order to obtain a best-fit solution to an infrared reflectance spectrum. Thus, the results of these two different experiments may be manifestations of the same physical imperfections at the interface. In one case, the imperfections may lead to optical changes, and, in the other, to the trapping of hydrogen. Furthermore, since the hydrogen density can be quite high at the interface, unwanted optoelectronic effects may occur in devices based on such processing techniques.

ACKNOWLEDGEMENTS

The authors wish to thank W.D. Laidig for the superlattice used in this study.

REFERENCES

[1] Garmire, E., Stoll, H., Yariv, A. and R. G. Hunsperger, Appl. Phys. Lett. 21 (1972) 87

[2] Scrifes, D. R., Streifer, W. and Burnham, R. D., IEEE J. Quant. Electon. QE-15 (1979) 917

[3] O'Hara, S., Speight, J. D., Leigh, P., McIntyre, N., Cooper, K. and O'Sullivan, P., Proc. Int. Conf. on Millimetric Waveguide Systems (London:IEE, 1976)

[4] Lindley, W. T., Phelan, R. J., Wolfe, C. M. and Foyt, A. G., Appl. Phys. Lett. 14 (1969) 197

[5] Liou, L. L., Spitzer, W. G., Zavada J. M. and Jenkinson, H. A., J. Appl. Phys. 59 (1986) 1936

[6] Chevallier, J., Dautremont-Smith, W. C., Tu, C. W. and Pearton, S. J., Appl. Phys. Lett. 47 (1985) 108

[7] Laidig, W. D., Blanks, D. K. and Schetzina, J. F., J. Appl. Phys. 56 (1984) 1791

[8] Wilson, R. G., Betts, D. A., Sadana, D. K., Zavada, J. M. and Hunsperger, R. G., J. Appl. Phys. 57 (1985) 5006

[9] Zavada, J. M., Hubler, G. K., Jenkinson, H. A. and Laidig, W. D., Proc. Mat. Res. Soc., Materials for Infrared Detectors and Sources (1987)

MULTIPOLAR PLASMA TREATMENTS OF $In_{0.53}Ga_{0.47}As$ SURFACE FOR MIS DEVICES APPLICATION

M. Renaud, P. Boher, J. Barrier, J. Schneider, J.P. Chanê

Laboratoires d'Electronique et de Physique Appliquée,*
3, Avenue Descartes - 94451 LIMEIL BREVANNES CEDEX (France)

Multipolar plasma treatments combined with a high vacuum system and a monitoring by ellipsometry have been developed to achieve high electrical quality InGaAs/Si_3N_4 interfaces. Interface states density in the low 10^{11} cm^{-2} eV^{-1} are obtained under controlled conditions, which is therefore suitable for a MISFET technology.

1. INTRODUCTION

1.1. InGaAs is now widely used for the fabrication of optoelectronic devices for long wavelength optical fibers telecommunications. Due to its excellent electronic properties and to the possibility to achieve the monolithic integration of optoelectronic devices with electronic components, the fabrication of InGaAs field effect transistors (FETs) is of great interest. Such FETs would also be suitable for high speed logic. Several different kinds of InGaAs FETs are considered and insulated gate FETs (MISFET) [1] and junction FETs (JFET) [2]are mainly investigated at present. Both devices require an insulator deposition step which is very critical. In the MISFET, the quality of the insulator/InGaAs interface directly determines the performances of the device while for Junction FETs (and also PIN photodiodes) the quality of the insulator/InGaAs interface settles the efficiency of the passivation and so the leakage current of these devices.
The achievement of a good insulator/semiconductor interface is difficult on III-V compounds because of their sensivity to high temperature and energetic particules. Besides, these compounds are not stable at the contact with their oxides [3] and native oxide is often thought to be responsible for the current drift observed in InP [4] and InGaAs [5] MISFETs.

Removal of native oxide prior to the deposition of an insulating layer should overcome these problems. However this step must be carefully realized, otherwise it can produce large damage on the surface. A process using a ultra high vacuum system and multipolar plasma treatments has been developed and successfully used on GaAs for the passivation of MESFETs [6], [7]. A similar approach is applied on InGaAs for the fabrication of MIS devices [8]. This paper will describe the various steps of the passivation scheme for InGaAs (oxide removal, nitridation and Si_3N_4 deposition). Dependence of treatment conditions on the electrical characteristics of MIS structures will also be discussed.

2. EXPERIMENTAL

The multipole first used by Limpaecher and

*LEP : A Member of the Philips Research Organization.

This work is supported by the European Community (ESPRIT project 927)

McKenzie [9] consists in an electron emitter (a hot filament) negatively biased (typically 20 to 100 V) with respect to the walls of a "magnetic container". The primary electrons emitted by the filament are confined by permanent magnets mounted around the walls of the chamber, and ionize the low pressure gases. This configuration produces chemically active species without high energy ions and dense plasmas at low pressures. The plasma conditions are dependent upon four parameters: the filament bias, the discharge current between filament and vessel, the total pressure and the composition of the gases in the plasma. Our experimental system, described elsewhere [10], is equipped with an ellipsometer to monitor the whole passivation process by kinetic (KE) and Spectroscopic (SE) Ellipsometry measurements.

In $_{0.53}Ga_{0.47}As$ layers lattice-matched to n^+ InP substrates are grown by vapor phase epitaxy using the chloride method. The InGaAs layers, about 3 μm thick, are unintentionally doped with a residual carrier concentration around 10^{15} cm^{-3} (n type), room temperature and 77 K mobilities respectively of the order of 9000 $cm^2\ V^{-1}s^{-1}$ and 40 000 $cm^2\ V^{-1}s^{-1}$. The samples are chemically cleaned in solvents and deoxidized in diluted HF before introduction into the ultra-high vacuum system. After passivation, a one-hour annealing at 400°C is performed under Ar or H_2. Finally, Ti/Au dots are evaporated through a metal mask and annealed 15 min at 400 °C under argon or hydrogen. Electrical characterizations are performed with a HP 4192 A impedance analyser to measure capacitance and conductance versus voltage from 500 Hz to 10 MHz.

3. PLASMA TREATMENTS OF InGaAs SURFACES

3.1. In situ removal of native oxide.

Removal of native oxide must not degrade the InGaAs surface by introducing roughness or perturbations in the stoichiometry. This leads to use plasma densities lower than 10^{11} cm^{-3} [8]. The removal of native oxide is achieved by a heating at 240 °C followed by a Hydrogen reducing plasma at 185 °C. Between each step, the sample is cooled back to room temperature in order to perform SE measurements required for a precise analysis of the surface state. Figure 1 shows the experimental KE trajectories measured at 3.7 eV during heating at 240 °C and further cooling under high vacuum. Theoretical trajectories corresponding to various native oxide thicknesses and substrate temperatures [8] are also drawn on this Figure and show that the initial surface is covered with about 10 Å of native oxide which is reduced to about 5 Å after the first heating. In spite of the high vacuum, the surface is slightly reoxidized during the first cooling (24 h). However, the second heating (185 °C) regenerates the previous surface. The H_2 reducing plasma at 185 °C during 1 min 30 s completely removes the native oxide but leaves also a reactive surface which reoxidizes during the second cooling.

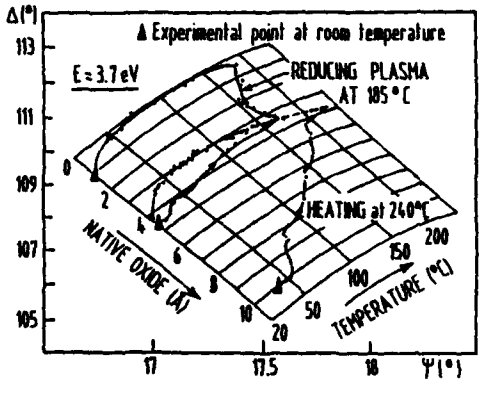


FIGURE 1

Kinetic Ellipsometry trajectories measured at 3.7 eV on InGaAs sample during in situ native oxide removal.

3.2. Native nitridation

As previously discussed, the clean surface is very reactive and needs to be stabilized. For this purpose, Nitrogen, which has the same valence as As, is used to saturate all

dangling bonds at the surface. After a heating at 205 °C which eliminates all adsorbed atoms, a N_2 plasma is performed at 150 °C during 1 min. KE trajectories represented on Figure 2 show that about 11 Å of native nitride are formed during this step.

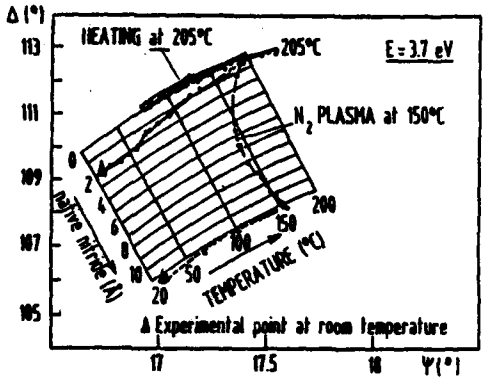


FIGURE 2

Kinetic Ellipsometry trajectories measured at 3.7 eV on InGaAs sample during the native nitridation.

3.3. Silicon nitride deposition

After the "cleaning" treatment of the surface, the sample is transferred under ultra high vacuum into the deposition chamber. A silicon nitride film, about 700 Å thick is deposited at room temperature on the passivation surface using a multipolar plasma of SiH_4 and N_2. The stoechiometry of this film has been shown to strongly depend upon the deposition conditions [11] and especially on the SiH_4/N_2 ratio, the discharge current, the total pressure and the substrate temperature for a given bias on the filament. Using spectroscopic ellipsometry data, films have been modelled by a mixture of pure Si_3N_4 and SiO_2. The concentration of oxygen in the film increases when the deposition rate increases. Besides, infra-red absorption analysis of films deposited on low doped Silicon wafers show that these films never contain Si-H while a few N-H exist in case of a deposition rate higher than 24 Å/min (Figure 3). SiN maximum absorption peak varies also with the deposition rate in agreement with the variation of stoechiometry.

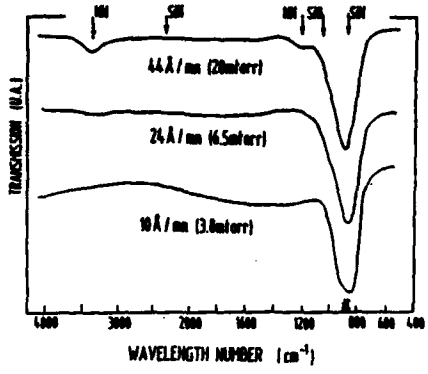


FIGURE 3

Infra-red absorption spectra of silicon nitride films.

4. ELECTRICAL CHARACTERISTICS

MIS structures have been first fabricated on untreated InGaAs surfaces (i.e. with native oxide) in order to test the influence of the stoichiometry of the silicon nitride film on the electrical properties of the interface. The 1 MHz C(V) curves presented on Figure 4 show that the modulation of capacitance increases when the deposition rate decreases or when the films approach stoichiometry. On the contrary, films containing a large amount of oxygen and NH bonds lead to poor electrical properties and a nearly pinned Fermi level. However the interface state density is still large (in the range of 10^{12} cm^{-2} eV^{-1}) in the best case.

Kinetic and spectroscopic ellipsometry have shown that multipolar plasma treatments of the InGaAs surface can completely remove the native oxide without optical degradation. Electrical characteristics of MIS structures fabricated using different plasma conditions for the oxide removal have been measured by the conductance method [12]. A typical set of C(V) curves recorded at various frequencies on an undegraded surface is shown in Figure 5. A large capacitance modulation is observed on the high frequency

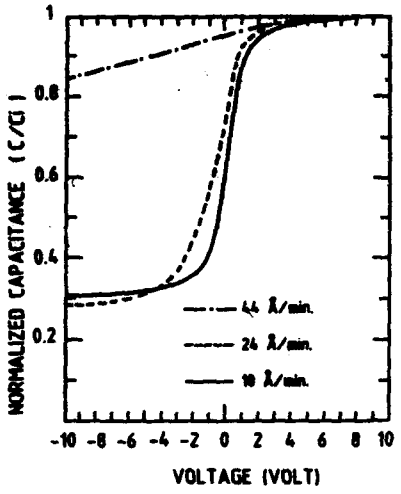


FIGURE 4

1 MHz capacitance normalized to the insulator capacitance Ci versus voltage of InGaAs MIS structures.

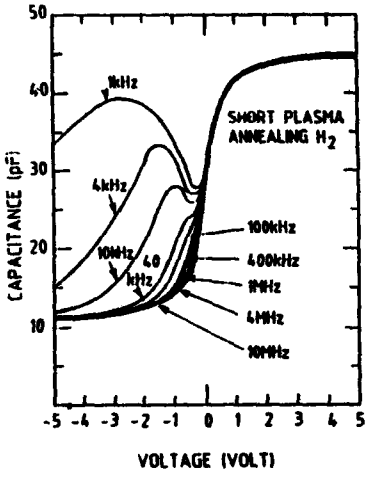


FIGURE 5

Typical set of C(V) curves at various frequencies for an InGaAs surface prepared in a plasma with a low concentration of ionized hydrogen.

10 MHz curve together with a small capacitance dispersion in the accumulation region. The 1 kHz C(V) curve clearly shows inversion of the InGaAs surface. For clarity, hysteresis has not been drawn but a clockwise hysteresis of about 0.7 V amplitude is observed with a sweep voltage ramp of 40 mV/s. This hysteresis can be explained by electron injection into the interface as proposed by Heiman and Warfield [13]. The interface state density N_{SS} was measured as a function of band bending on MIS structures by the conductance technique.

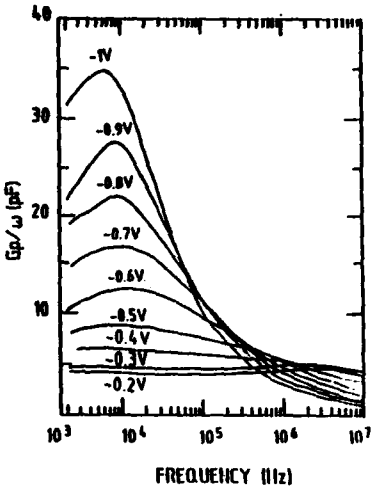


FIGURE 6

Equivalent parallel conductance (Gp/w) versus frequency of MIS structure described in figure 5.

In Figure 6, the equivalent parallel conductance divided by the angular frequency (Gp/w) is represented versus frequency for a MIS structure fabricated on a GaInAs surface treated in a plasma containing a low concentration of ionized hydrogen. The Si_3N_4 film has been deposited under optimized conditions discussed previously and annealed under H_2. These Gp/w versus frequency curves follow the classical curves for the Si-SiO_2 system (but at higher frequencies due to the smaller band gap of InGaAs), being nearly flat near flatband and peaked in weak inversion. Figure 7 summarizes N_{SS} data determined by the conductance technique on MIS structures in which the plasma conditions for oxide removal and the annealing gas are the only varied parameters. A low concentration of ionized hydrogen leads to a low

interface state density in the range of $(3\text{-}5)\times10^{11}cm^{-2}eV^{-1}$. Nearly pinned Fermi level is obtained when increasing the exposure time and the density of the plasma treatment. Annealing under hydrogen is also influencing the electrical properties of the Si_3N_4/InGaAs interface.

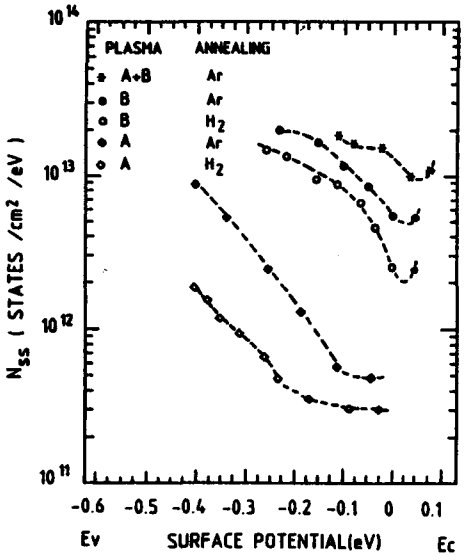


FIGURE 7

Distributions of the interface state density in the band gap of InGaAs for various oxide removal treatments (A : low concentration of ionized hydrogen, B : high concentration of ionized hydrogen).

CONCLUSION

A complete passivation scheme has been investigated on InGaAs surfaces. Combining in-situ kinetic and spectroscopic ellipsometry measurements, a range of experimental conditions in which no optical degradation occurs on the surface has been determined. Electrical characterizations of the Si_3N_4/InGaAs interface on MIS structures have shown that optimum conditions are even more restricted but that excellent interface properties can be obtained using multipolar plasma treatments in a ultra high vacuum system. Stoichiometry of the Si_3N_4 film itself has also been shown to strongly affect the electrical properties of MIS structures. N_{SS} in the low 10^{11} cm^{-2} eV^{-1} indicates that the optimized passivation process is suitable for a MISFET gate insulator.

ACKNOWLEDGEMENTS

The authors would like to thank Y. Hily for multipolar plasma treatments of InGaAs surface and E. Boucherez for the fabrication of test structures.

REFERENCES

[1] P.D. Gardner, S.Y. Narayan and Y.H. Yun, Thin Solid Films, 117 (1984) 173-190.

[2] D. Wake, A.W. Nelson, S. Cole, S. Wong, I.D. Henning and E.G. Scott, IEEE Elect. Device Letters, Vol. EDL 6 N° 12 (1985) 626.

[3] G.P.Schwartz, B. Schwartz,J.E. Griffiths and T. Sugano, J. Electrochem. Soc. 127 (1980) 2269

[4] K.M. Geib, S.M. Goodnick, D.Y. Lin, R.G. Gann, C.W. Wilmsen and J.F. Wager, J. Vac. Sci. Technol. B2, 516 (1984).

[5] M. Taillepied, S. Gourrier, Applied Physics Letters 48 (1986), 978.

[6] J.B. Theeten, S. Gourrier, P. Friedel, M. Taillepied, D. Arnoult and D. Benarroche, Mat. Res. Soc. Symp. Proc. Vol. 38 (1985) 499.

[7] P. Boher, F. Pasqualini, J. Schneider, Y. Hily, 6th Int. Coll. Plasmas and Sputt., Antibes, June 1st to 6th (1987).

[8] P. Boher, M. Renaud, J.M.Lopez-Villegas, J. Schneider and J.P. Chanê, INFOS 1987, Leuven, April 13th-15th (1987).

[9] R. Limpaecher, R.R. Mackenzie, Rev.Sci. Instrum. 44, 6, (1973).

[10] D. Arnoult, Thèse de Docteur Ingênieur, INSA de Lyon, France (1986).

[11] P. Boher, M. Renaud, L.J. Van Izendoorn, J. Barrier and Y. Hily, Submitted to Journal of Appl. Physics.

[12] E.H. Nicollian, A. Goetzberger, Bell Syst. Tech., 46 (1967) 1055.

[13] F.P. Heiman and G. Warfield, IEEE Trans. on ED N° 12 (1965) 167.

VISIBLE LIGHT a-SiC THIN FILM LED AND ITS APPLICATION TO NEW OE-FUNCTIONAL ELEMENTS

Yoshihiro HAMAKAWA, Dusit KRUANGAM, Hiroaki OKAMOTO and Hideyuki TAKAKURA

Faculty of Engineering Science, Osaka University,
Toyonaka, Osaka 560, Japan.

A visible-light injection-type electroluminescence thin film diode made of amorphous silicon carbide (a-SiC TFLED) has been developed. The device has a basic structure of p (a-SiC)/i (a-SiC)/n (a-SiC). The emission color could be controlled from red to green by adjusting the carbon content x in the a-$Si_{1-x}C_x$ luminescent i-layer. The brightness of 5 cd/m^2 was obtained for the yellow LED with a forward injection current density of 200 mA/cm^2. A series of technical data on the device fabrication technology, injection efficiency improvement and resulting device characteristics are presented, and discussed in relation to the carrier injection and recombination mechanisms. The developed a-SiC LED has some significant advantages over the conventional crystal LEDs, such as wide area, ease of fabricating integrated type multi-color or tunable color LEDs, and low cost. Utilizing these unique significances, new type of opto-electronic functional elements are proposed and discussed.

1. INTRODUCTION

Since the recent success of valency electron control in the glow discharge produced amorphous Silicon Carbide (a-SiC:H) in 1981 (1), so called amorphous silicon alloy age has opened up, and a group of new materials such as amorphous silicon-germanium (a-SiGe:H), amorphous silicon-nitride (a-SiN:H) and amorphous silicon-tin (a-SiSn:H) have been successively born in the following few years. The significance of this material innovation is that one can control electrical-, optical- and also opto-electronic-properties by controlling atomic compositions in the mixed alloys. Therefore, a wide varieties of application fields has also been opened up for this new electronic material. In fact, a-SiC/a-Si heterojunction solar cells (2), a-Si/a-SiGe stacked solar cells (3), superlattice devices (4), a-Si/a-SiN thin film transistors (5), photo-receptors (6), X-ray sensor (7), color sensors (8) etc have been developed and some of them are already in commercial market.

On the electroluminescence in a-Si, fastly, Pankove et al have reported an infrared emission from a shottky barrier interface of a-Si:H and also a-Si:H p-i-n junction at low temperature in 1976 (9). Since then, the same kind of observation has been reconfirmed by several groups (10-13). Recently, a visible electroluminescence has been observed by Munekata et al (14) and Matsunami et al (15) in ac-driven a-SiC:H layer sandwitched with insulating layers. The operation voltage was more than 100 V, because this emission was based upon an intrinsic electroluminescence. For the practical application to a thin film display panel and/or opto-electronic coupled functional element, we do need more low threshold voltage device by an injection type electroluminescence.

We have challenged along this line and tried to develop a visible light injection electroluminescence a-SiC:H p-i-n diode (a-SiC TFLED), and succeeded in obtaining red, orange, yellow and green color emissions.

The purpose of this paper is to present technical data on the material preparation and its optoelectronic properties, cell structure, LED characteristics and mechanism identifications. The developed TFLED has various attractive advantages; it can be

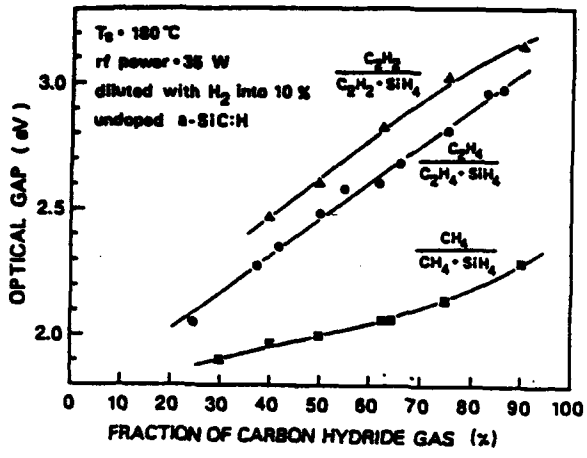


Fig.1 Variations of optical energy gap of a-SiC:H with carbon hydride gas fraction in the source gas SiH_4.

operated with a much lower applied voltage as compared with an intrinsic EL type, and this results in the possibility of coupling with its driving circuits, e.g., crystalline integrated circuit or a-Si thin film film transistor, and one can construct a 3-dimensional opto-electronic integrated circuit (OE-IC). In this paper, possibilities of this kind of new applications are also proposed and discussed.

2. PREPARATIONS AND PROPERTIES OF a-SiC:H

a-SiC:H films were prepared mostly by the cross field plasma CVD system (16). The plasma is excited by rf (13.56MHz) power in parallel to the substrate surface. An additional DC bias is applied vertically to the substrate. As is now well recognized, material properties are sensitive to the electric potential near the substrate surface. The DC bias is then used to adjust this substrate potential so as to obtain the required material properties, such as high electrical conductivity and wide optical band gap of injector layers. The substrate temperature is 180°C. The total gas pressure during the deposition is 1 torr. CH_4, C_2H_4, C_2H_2 or TMS (tetramethylsilane) were used as carbon source gases. Figure 1 shows the relation between the optical energy gap of undoped a-SiC:H and the fraction of carbon hydride gas used in the reaction gas. The optical gap of a-SiC:H increases monotonically from 1.7 eV to more than 3.0 eV with increase in the carbon gas fraction. It should be noted that the optical gaps of C_2H_4- and C_2H_2- based a-SiC:H are larger than those of CH_4- based a-SiC:H. Since the emissive photon energy should be essentially limited by the optical gap energy, a-SiC:H possessing a wide band gap is preferable for obtaining visible emission.

The a-SiC TFLED developed in this work has a structure of p (a-SiC)/i (a-SiC/n (a-SiC). In this type of device, p (a-SiC) and n (a-SiC) act as the injectors of holes and electrons into the luminescent active i-layer, and thereby the LED performance primarily strongly depends on the carrier injection efficiency. Therefore, it is of great importance to have a wide-gap with highly conductive p- and n- type a-SiC as injector layers. For this purpose, we adopted the cross-field plasma CVD technique. It has already been demonstrated (16) that an application of positive DC bias voltage at the substrate results in a drastic increase of the dark conductivity of boron doped a-Si:H. In this work, the similar effect of DC bias voltage has also been found in a-SiC:H systems. Moreover, an increase of the optical gap has been found for example, from E_o=1.87 eV to 2.0 eV (17).

Another approach we have tried to prepare a wide-gap highly-conductive p type a-SiC injector is by using Electron Cyclotron Resonance (ECR) plasma CVD. The details of the ECR CVD and preparation techniques have been reported in our recent paper (18). A result of the relation between the dark conductivity and the optical energy gap of boron doped p a-SiC prepared by ECR plasma CVD and by conventional rf plasma CVD has been shown in Fig.2. As can be seen, as the optical energy gap increases, the dark

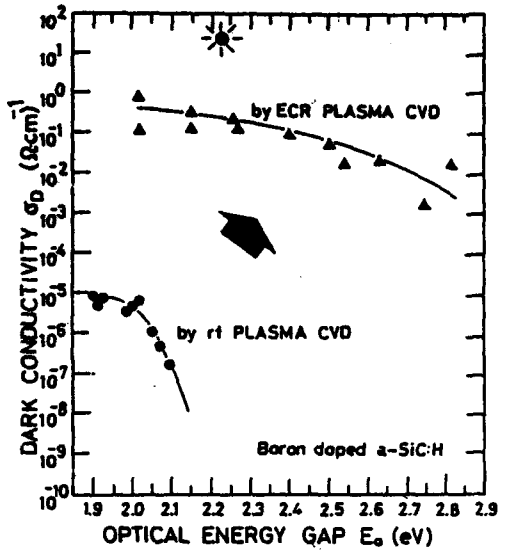


Fig. 2 Relation between the dark-conductivity and the optical gap E_0 of boron doped p-type a-SiC prepared by ECR plasma CVD and rf plasma CVD.

conductivity (σ_D) for the materials prepared by rf plasma CVD rapidly decreases. While σ_D for the materials prepared by ECR CVD still retains higher than 10^{-3} (S/cm) even the optical energy gap exceeds 2.5 eV. The highest σ_D obtained is 20 (S/cm) as shown in the mark-☼- and the material has an extremely small ΔE (E_F-E_V) less than 0.1 eV with a low content of hydrogen content around 5%. The results from Raman spectra measurement have shown that the material prepared by ECR CVD has a microcrystalline phase. It is noted here to avoid the degradation of ITO/SnO_2 electrode caused by the hydrogen radicals in the ECR plasma that the TFLED device structure has been modified to be an invert type, that is ITO/SnO_2/n-i-p/Al, in which the n- and i-layers were prepared by rf plasma CVD and p-layer by ECR CVD.

3. BASIC CHARACTERISTICS OF a-SiC p-i-n TFLED

a-SiC TFLED developed has a junction structure of glass/ITO/SnO_2/p (a-SiC)/i (a-SiC)/ n (a-SiC)/Al as shown in Fig.3. The thicknesses of p- and n- (a-SiC) injector layers are 150 A and 300 A, respectively. The optical band gap of the luminescent i-layer was varied in the range from 2.2 to 3.7 eV. The substrate temperature and the total gas pressure during the deposition are 180°C and 1 torr, respectively. It should be noted that the a-SiC TFLED formed on a milky-rough ITO/SnO_2 exhibits larger electroluminescent (EL) intensity than that formed on a smooth ITO/SnO_2 by one order of magnitude. This may be due to both a better electrical contact between ITO/SnO_2 and p a-SiC, and a reduction of apparent refractive index of the p-layer.

Figure 4 shows a schematic band diagram of a

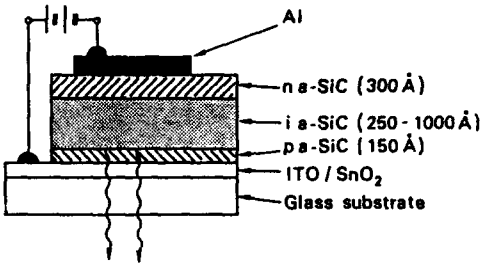


Fig. 3 A schematic illustration of a-SiC p-i-n junction TFLED formed on a glass/ITO/SnO_2 substrate.

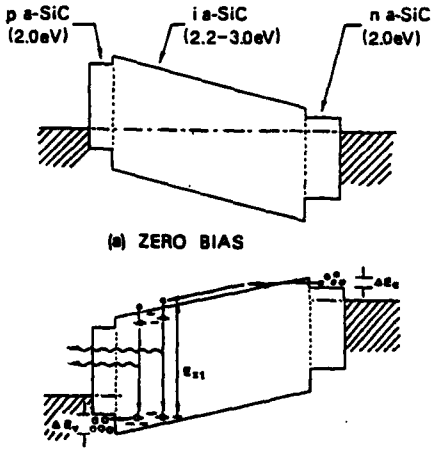


Fig. 4 Energy band diagrams of a-SiC p-i-n junction TFLED at thermal equilibrium (a) and forward bias (b) conditions.

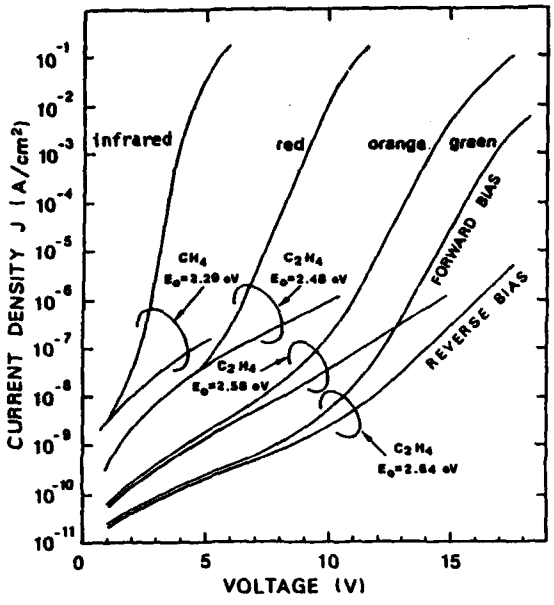


Fig.5 J-V characteristics of a-SiC p-i-n TFLED. E_0 denotes the optical gap of i-layer. The thickness of i-layer in the case is 1000A.

SiC p-i-n junction in thermal equilibrium (a) and forward bias (b) conditions. In this LED, injection electroluminescence can be observed only in a forward bias condition. In order to obtain a visible luminescence, the optical band gap of the i-layer has to be larger than about 2.5 eV, as deduced from the results of the photoluminescence. While the optical band gaps of p- and n- a-SiC layers should be chosen around 2.0 eV to ensure the effective valency controllability to p or n. Thus, there should exist band discontinuities at the p/i (ΔE_v) and i/n (ΔE_c) interfaces. These band discontinuities have been estimated from the results of the internal photoemission measurement (19). For instance, for E_o of i layer = 2.58 eV and E_o of p-,n- layer = 2.00 eV, the conduction band discontinuity ΔE_c at the i/n interface and the valence band discontinuity ΔE_v at the p/i interface are 0.19 eV and 0.48 eV, respectivaly.

Figure 5 shows logarithmic plots of the J-V characteristics measured at room temperature for various optical gaps of i layers(2.29, 2.48, 2.58 and 2.64 eV.). In these samples, the thickness of the i layer was set at 1000 A. As can be seen, even though a-SiC is used for the i layer, the rectification ratio of more than 10^6 is obtained at 5 V for the sample with E_o=2.29 eV. However, it should be remarked that when the optical gap, or in other words the carbon content, of the i layer increases, the threshold voltage tends to increase. The threshold voltage is here defined as the voltage required to give a current of 10^{-7} A/cm^2. This might be partly due to the increase of series resistance in the i layer itself and to the increase of the barrier height at both p/i and i/n interfaces.

Here, we will briefly discuss the carrier transport mechanism in the a-SiC p-i-n junction. When the tunneling process dominates the diode current in the high voltage region, it is expressed by (20):

$$J \propto E^2 \exp[-4\sqrt{2m^*}(\phi_B)^{\frac{3}{2}}/3q\hbar E], \qquad (1)$$

where J is the current density, E the applied electric field across the i layer, and m^* denotes the effective mass of the carrier, q the electron charge, $\hbar$ the Planck constant. Φ_B is the effective barrier height for tunneling which in this case corresponds to the band discontinuities ΔE_c or ΔE_v.

Figure 6 shows J-V characteristics of a-SiC p-i-n junctions for forward bias voltage, where the optical gap of the i layer is 2.68 eV and the i layer thickness is varied from 500 to 750 A. If a tunneling process dominates and an applied forward voltage V is uniformly distributed through the sample of thickness d, eq.(1) suggests that $\log(J/V^2)$ should have a linear relation with 1/V. In order to verify this relation, the J-V characteristics are replotted into the $\log(J/V^2)$ vs 1/V in the inset of Fig.6. The inset figure reveals that $\log(J/V^2)$ is in proportion to 1/V in the high

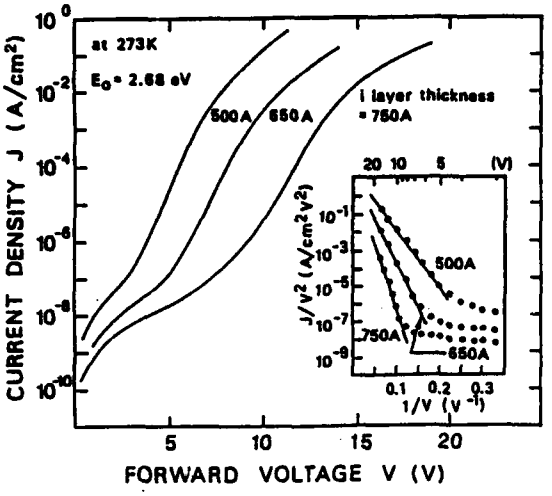


Fig.6 J-V characteristics of ORANGE a-SiC TFLED as a function of i-layer thickness. The optical gap of i-layer is 2.68 eV. The inset shows J/V^2-$1/V$ plots indicating a tunneling injection in the TFLEDs.

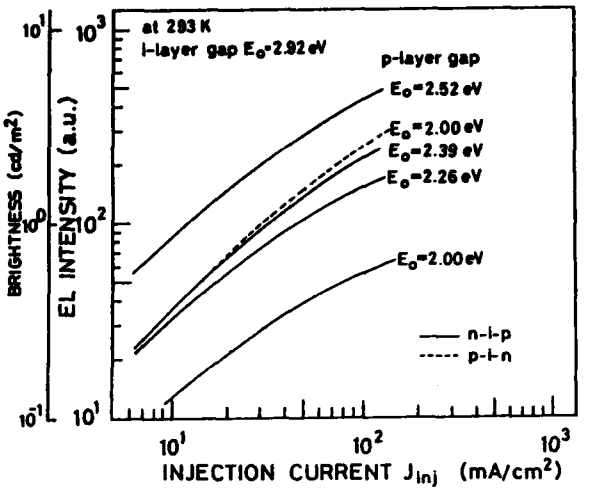


Fig.7 Log-log plots of EL intensity vs injection current density for various optical gaps of p-layers.

voltage region, where we can observe the electroluminescence. The effective barrier height can be evaluated from the slope of these curves. Assuming that the effective carrier mass is just equal to the free electron mass, the effective barrier height has been found to be about 0.25±0.03 eV. This value is close to the conduction band discontinuity at the n/i interface ΔE_c=0.22 eV, which was estimated from the internal-photoemission. The above discussions lead to a conclusion that the dominant current in the high voltage region (5-10 V) is the electron tunneling current through the energy barrier at the i/n interface. Since the energy barrier at the p/i interface ΔE_v is larger than ΔE_c, the hole tunneling current is considered to make only a minor contribution to the total diode current. For the electroluminescence, however, not only electrons but also holes must be injected into the i-layer. Therefore, the hole tunneling current will play a central role in the electroluminescence (EL) property. Fig.7 shows the relation between the EL intensity and injection current density measured at room temperature. The solids lines are for the n-i-p types, of which the p-layers were prepared by ECR CVD. For comparison, the result of p-i-n type, in which all layers were prepared by rf plasma CVD is shown with the broken line. The optical energy gaps for n- and i-layers were 2.0 eV and 2.92 eV, respectively. It is likely that the n-i-p type is somewhat inferior to the p-i-n type. This might be due to: 1) the effect of residual phousphorus impurity inclusion in the i-layer (21) and 2) the effect of internal-absorption since the radiative recombination mainly takes place in the i-layer near the i/p interface(17). However, increasing the optical energy gap of the p-layer in the n-i-p type, the EL intensity gradually increases and becomes higher than that of the reference p-i-n type by two times of magnitude. The EL spectra for n-i-p LEDs are shown in Fig.8. As the optical energy gap of p-layer increases, the EL spectrum shifts towards shorter wavelength and the emitting color drastically changes from red to orange and to yellow. The

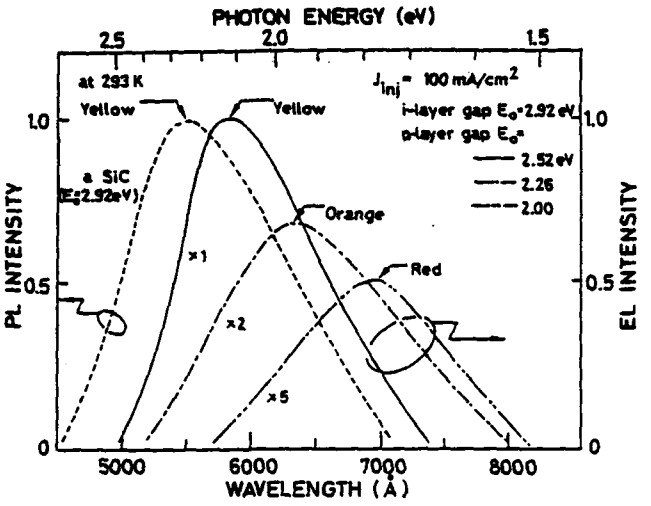


Fig.8 EL spectra of a-SiC TFLED as a function of the optical gap of p-layer. The spectrum of the i-layer used in the TFLED is also shown for comparison. The thickness of i-layer in the case is 500A.

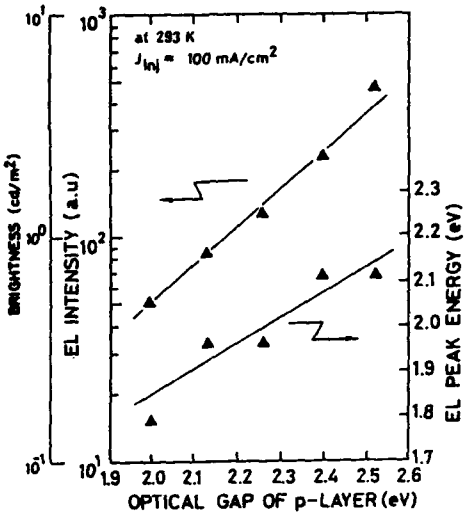


Fig.9 Summarized data on the dependence of EL intensity, EL peak energy on the optical gap of p-layer.

dash line shown in the figure is the PL spectrum (excited by 365nm light) of i a-SiC film identical to the one used in the n-i-p junctions. It should be noted that as the optical energy gap of the p-layer increases, the EL spectra shift closer to the position of the PL spectrum. Fig.9 summarizes the dependence of the EL intensity and the peak energy of the spectrum on the optical energy gap of the p-layer. The results in Figs.8 and 9 indicate that the EL performances strongly depend on the properties of the p-type injector. The increase of the optical energy gap of the p-layer decreases the barrier height for holes to tunnel, or in other words, lifts holes up to be injected into the localized tail states locating closer to the valence band edge of the i-layer and this results in the improvement of the EL performances. Figure 10 shows an example of the application of a-SiC LED as a large area display. The emitting area is about 115 mm^2.

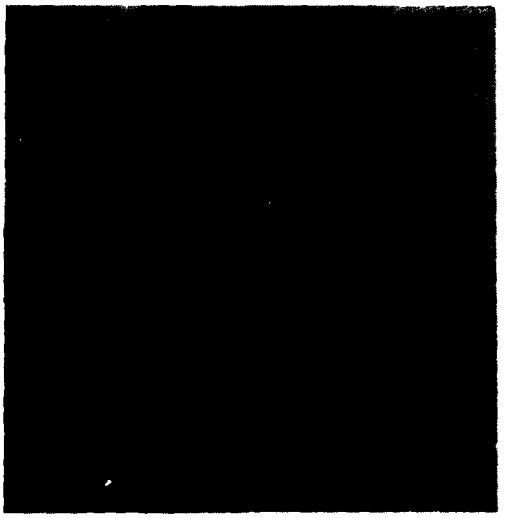


Fig.10 Example of a real yellow emitting a-SiC TFLED. The area of the rabbit is about 115 mm^2.

The color is yellow and the brightness is about 5 cd/m^2 at the injection current of 20mA and about 10V.

At present, further efforts have also been being done to increase the brightness by another one order of magnitude in our laboratory. By combining a new preparation technology such as ECR microcrystalline SiC, with new device structures; e.g., superlattice i-layer and stacked p-i-n/p-i-n etc.(22), we hope that the final goal of the brightness for a practical use more than 50 cd/m^2 should be overcome in the near future.

4. APPLICATIONS TO OPTOELECTRONIC FUNCTIONAL ELEMENTS

Another realistic application of TFLED at present stage of brightness would be opto-electronic (OE) functional elements. Since first proposal of optoelectronic devices by E.E. Loebner in 1959 (23), considerable efforts have been paid to develop various OE functional elements such as image convertor, image storage devices etc (24). However no practical devices have been developed so far due to mismatching of opto-electronic coupling circuit and also technical difficulties of the integration of electroluminor (EL panel) and photosensor (II-VI compound photoconductor). We have been conducting a series of basic research on a new optoelectronic integrated circuit (OE-IC) with full utilization of unique advantages of a-SiC TFLED and a-Si solar cell technologies such as a wide area with low driving voltage and ease of multi-layered structure by low temperature process on any foreign substrates.

Figure 11(a) shows a construction of the basic three OE-logic circuits by combinations of TFLED and TFPDA (Thin Film Photo-Diode Array), and (b) is the corresponding equivalent OE circuits. By combining these basic logic circuits with the electrical or optical storage, one can easily construct an OE-Shift Registor, a Flip-Flop, a functional generator and an adder. Therefore an OE-computer might be developed.

With use of big significances of two dimentional information processing in this OE-IC, the image convertor, image storage and pattern recognition devices could be also developed. Fig. 12 shows a schematic illustration of the opto-electronic image translator. In the figure, DIPUs (Distributed Image Processing Unit) are constructed by integrated TFLEDs and a-Si solar cells. An extremely fast information processing can be done because of two dimentional parallel access. A series of systematic basic research has been just started and in progress.

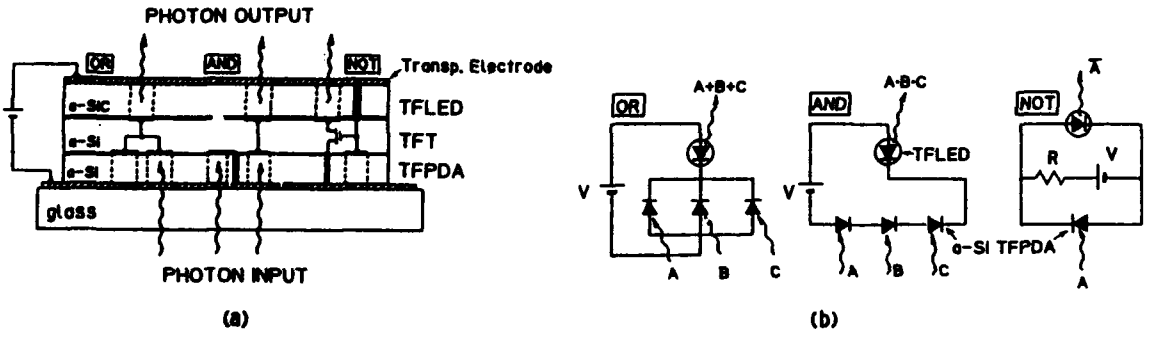


Fig.11 An example of OE-Logic Components constructed by TFLED, TFPDA and a-Si TFT (a) and their equivalent circuits of OR, AND and NOT functions (b).

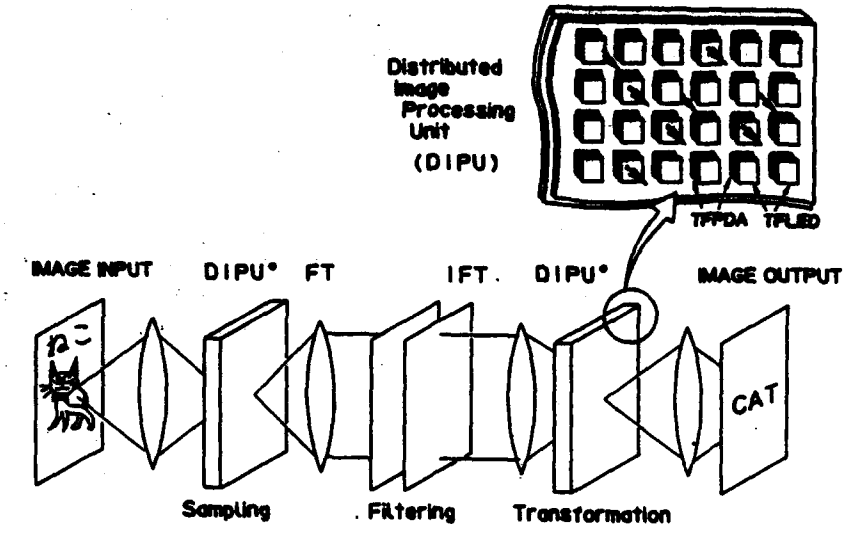


Fig.12 A proposed OR-IC device which has functions of not only image translation, pattern recognition but also image association which might be accomplished by combining with image characterization process.

REFERENCES

(1) Y. Hamakawa and Y. Tawada: Int. J. Solar Energy 1 (1982) 251.
(2) Y.Tawada, M.Kondo, H.Okamoto and Y.Hamakawa: Solar Energy Materials. 6 (1982) 299.
(3) S.R. Ovshinsky; MRS 1985 Spring Meeting, San Francisco, (1985) 251.
(4) S. Tsuda, H. Tarui, T. Matsuyama, H. Haku, K. Watanabe, Y. Nakashima, S. Nakano and Y. Kuwano: Proc.2nd Int. Photovoltaic Sicience and Engineering Conf., Beijing China,(1986) 409.
(5) For example : JARECT, edited by Y.Hamakawa (Ohm-Sha & North-Holland, Tokyo, Amsterdam, New York, Oxford, 1983) vol.6. p.252.
(6) I. Shimizu: J. Non-Crys. Solids 77&78 (1985) 1363.
(7) Wei Guang-Pu, H.Okamoto and Y.Hamakawa: Jpn.J.Appl.Phys. 24(1985)1105.
(8) S. Nakano, T. Fukatsu, M. Takeuchi, S. Nakajima and Y. Kuwano: Proc. 3rd Sensor Symposium, Tsukuba Japan (1983) 97.
(9) J.I. Pankove and D.E. Carson: Appl.Phys. Lett. 29 (1976) 620.
(10) T.S. Nashashibi, I.G. Austin, T.M. Searle, R.A. Gibson, W.E. Spear and P.G. LeComber : Phil.Mag.B45 (1982) 553.
(11) K.S. Lim, M. Konagai and K. Takahashi: Jpn.J.Appl.Phys. 21 (1982) L473.
(12) A.J. Rhodes, P.K. Bhat, I.G. Austin, T.M. Searle and R.A. Gibson: J.Non.-Cryst Solid. 59&60 (1983) 365.
(13) P.K. Bhat, T.M. Searle, I.G. Austin, R.A.Gibson and J.Allison: Solid State Comm. 45 (1983) 481.
(14) H. Munekata, H. Kukimoto: Appl.Phys.Lett. 42(5) (1983) 432.
(15) H. Matsunami, M. Yoshimoto, Y. Fujii and J. Saraie: J.Non-Cryst.Solids 59&60 (1983) 569.
(16) S. Hotta, N. Nishimoto, Y. Yawada, H. Okamoto and Y. Hamakawa: Jpn.J.App.Phys 21 suppl.21-1 (1982) 289.
(17) D. Kruangam, T. Endo, M. Deguchi, W. Guang-Pu, H. Okamoto and Y. Hamakawa, OPTOELECTRONICS -Devices and Technologies (Mita Press) 1 (1) (1986) 67.
(18) Y. Hattori, D. Kruangam, K. Katoh, Y. Nitta, H. Okamoto and Y. Hamakawa, 19th IEEE Photovoltaic Specialists Conf., New Orleans, U.S.A., May, 1987.
(19) D. Kruangam, M. Deguchi, T.Endo, H.Okamoto and Y.Hamakawa: to be submitted.
(20) S.M. Sze: Physics of Semiconductor Devices (John Wiley & Sons, New York, 1981) 2nd.,p.402.
(21) K. Lim, M. Konagai and K. Takahashi, Jpn. J. Appl. Phys, 21 (1982) L473.
(22) D. Kruangam, M. Deguchi, T.Endo, W. Guang-Pu, H. Okamoto and Y. Hamakawa, Extended Abstract of the 18th 1986 Int. Conf. Solid State Devices and Materials, Tokyo (1986) 683.
(23) E.E. Loebner: "Solid State Optoelectronics", RCA Rev, 20 (1959) 715.
(24) S. Larach: "Photoelectronic Materials and Devices (D. Van Nostrand Company, Inc. New York, London, 1965) 295.

Session A1.3

MOS Reliability
I

Chairman: T.F. Retajczyk

Monday, September 14, 1987

COMPARISON BETWEEN HOT-CARRIER DRIFT AND RADIATION DAMAGE IN MOS DEVICES

Anant G. Sabnis

AT&T Bell Laboratories
1247 S. Cedar Crest Boulevard, Allentown, PA 18103 U.S.A.

Damages due to hot-carrier injection and exposure to ionizing radiation are linked through the kinetics of hole trapping and their conversion to fast states at the Si/SiO_2 interface. The processing conditions and the presence of hydrogen ambient have similar impact on the device response to these two degradation mechanisms. For a given probability of hot hole generation and their injection, an increased radiation softness of an MOS device would result in an increased susceptibility to hot-carrier degradation.

1. INTRODUCTION

The degradation of MOS IC's due to injection of hot-carriers is no longer a scientific curiosity. This mechanism is likely to become a major stumbling block in the advance of the silicon IC technology. Accordingly, the study of this mechanism has become an integral part of the technology development. It has been now well recognized that, as a result of hot-carrier injection, structural and chemical changes occur at the Si/SiO_2 interface. The changes manifest in terms of trapped carriers and interface states in short-channel length MOSFETs. Several physical models that describe these changes have been advanced in the literature [1-4].

Based on the characterization of devices over a wide range of operation, we have proposed a physical model for device degradation which invokes the injection of hot holes and their conversion to interface states [4]. The exposure to ionizing radiation is known to trap holes at the Si/SiO_2 interface. We have observed that the two mechanisms, the hot hole injection and the ionizing radiation, involve similar kinetics of hole trapping and conversion of holes into fast states. As a result, the impact of these mechanisms on device characteristics are similar. The purpose of this paper is to discuss those similarities. The terms interface states and fast states are used interchangeably.

2. EXPERIMENTAL

In these experiments, MOSFETs fabricated with the n-channel Si-gate technology were used. Some devices had a final encapsulation layer of plasma deposited silicon nitride (SiN-Cap) and some devices had no encapsulation (No-Cap). For radiation damage studies, large area (200μm x 200μm) devices and a laboratory Co^{60} source of gamma rays were used. For hot-carrier damage studies, short-channel length MOSFETs were electrically stressed under dc and ac conditions [4, 5]. The devices were characterized, before and after damage and after anneal, over a wide range of operation in the forward and reverse modes (source and drain interchanged). Only those devices which had initially the same forward and reverse mode characteristics (symmetrical devices) were selected for these experiments.

In addition, some DRAMs were irradiated under bias and others were operated under electrical stress for hot-carrier damage. Drift in their V_{DDmin} is measured as a function of radiation dose or stress time. V_{DDmin} is the minimum voltage required for a DRAM to operate under a given set of timing conditions.

The MOSFET data is analyzed, first by calculating the difference in the gate voltage (ΔV_G or ΔV_{GS}) required to pass the same channel current

(I_{DS}) for the same V_{DS} and V_{BB} before and after damage, and then by plotting the difference as a function of log I_{DS}. The drift at I_{DS}=10^{-11} Amp gives the magnitude of trapped charges (ΔQ_t) and the slope of the curve gives the density of fast states (ΔD_{it}) at the Si/SiO_2 interface. The Si/SiO_2 interface will be simply referred to as "the interface" in the following text. The details of the characterization are given in the references [4-6] cited at the end of the text.

3. RESULTS AND DISCUSSIONS

Drifts in transfer characteristics of a MOSFET due to hot carriers are shown in Figure 1. The curves A and C are the (semilog) initial characteristics, both in forward and reverse modes, taken at V_{DS} = 5V (saturation region) and V_{DS} = 0.1V (triode region) respectively. The device was symmetrical before aging. After aging, the reverse mode characteristic at V_{DS} = 5V separated as shown by the dashed curve B, and the device is no longer symmetrical. At V_{DS} = 0.1V, the device remains symmetrical even after aging and the drift is shown by the dashed curve D. The triode region characteristics at V_{DS} = 0.1V are also plotted in the linear mode; after aging, the initial curve E drifts to F. The linear mode characteristics show only the transconductance degradation with a negligible change in the extrapolated threshold voltage (V_T). From these characteristics alone it may appear that the observed drift is caused by an increase in the drain resistance due to electrons trapped in the oxide above the drain diffusion [3] or due to the creation of fast states at the interface near the drain junction. However, when the aging was continued, we observed negative drift in the forward mode characteristics at V_{DS} = 5V, spreading of the fast states into sub-threshold region and a positive drift in the linear mode V_T as shown in Figure 2. From the linear mode characteristics alone, the positive ΔV_T may appear to be due to trapping of electrons at the interface near the drain. However, the other two observations listed above and an additional observation of channel shortening effect in the forward mode lead us to believe that this drift is due to injection of holes into the oxide [4]. The injected holes are trapped at the interface near the drain region. Some of these trapped holes convert to fast-states by capturing electrons or negative ions such as H^-. Further detailed experiments [5] indicate that there is a lifetime (τ_c) associated with these trapped holes after which they convert to neutral acceptor-like fast states. This process of conversion is a characteristic of the interface, and it is through this process that we observe a link between the hot-hole damage and the radiation damage.

In an ionizing radiation (Co^{60}) environment, holes are trapped at the interface independent of the direction of the electric field at the interface. When the gate is negative with respect to the substrate, the holes are injected from silicon into the oxide [6]. This situation is analogous to the hot-hole injection. With positive gate voltage, it is generally accepted

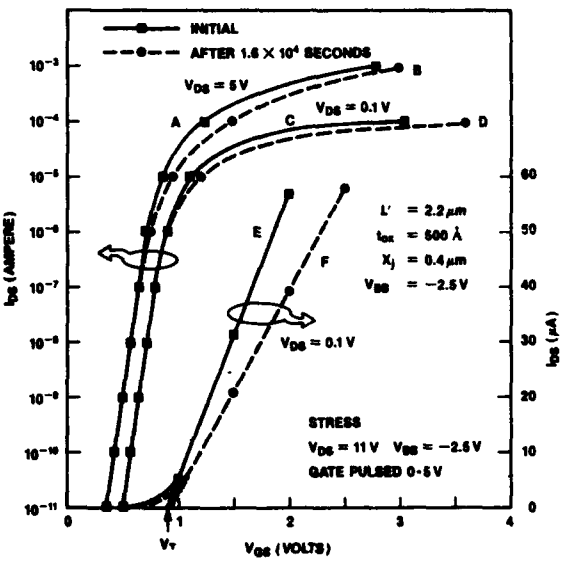


FIGURE 1: *Gain degradation due to hot-carrier injection.*

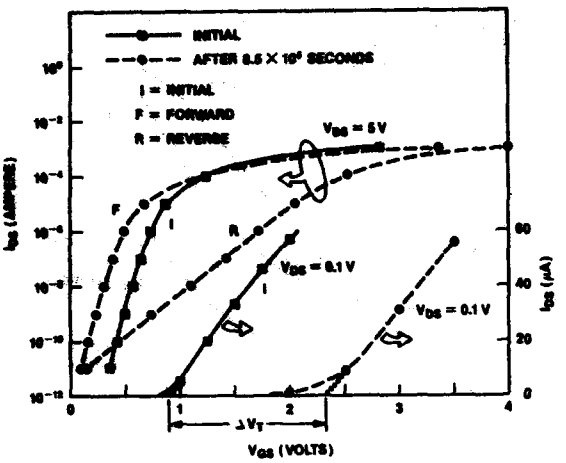


FIGURE 2: ***Degradation of the device of Figure 1 after 236 hours of stress.***

that the holes generated in the oxide are trapped at the interface. Large area MOSFETs (200μm x 200μm) were exposed to Co^{60} source of gamma rays under various biases and their transfer characteristics were measured in the forward and reverse modes. The curve A in Figure 3 is the drift under positive bias. The curve is almost parallel to the log I_{DS} axis suggesting that, at the interface, holes are trapped uniformly and a few fast states are created in the high current region. On the other hand, under negative bias, a significantly large number of fast states is created in the high current region as shown by the curve A in Figure 4. The curve B in Figure 4

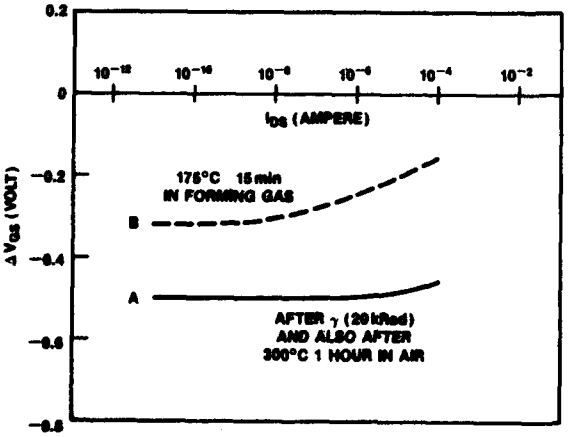


FIGURE 3: ***Saturation region (V_{DS} = 5 V) drift in 200μm x 200 μm MOSFET without encapsulation.***

is a typical drift due to hot-hole injection which differs from the curve A only by the magnitude of the trapped positive charges. Otherwise, the two curves are similar. A comparison of the curves A in Figures 3 and 4 suggests that, when the holes are injected from silicon into oxide, the fast states readily appear at the interface. When the device in Figure 3 was heated in air at 300 ° C for 1 hour, no change was observed in curve A. However, when the same device was heated at 175 ° C for 15 minutes in forming gas, a process of conversion of holes into fast states is observed, which is inferred from a reduced negative drift and a simultaneous increase in the slope of the curve B in Figure 3. This shows that the presence of hydrogen accelerates the conversion process or, in other words, reduces the lifetime of the holes. This is true even in the case of hot-hole injection phenomenon. In Figure 5, aging rates of no-cap devices in air and in hydrogen clearly show the impact of hydrogen on the build-up of fast states. The device with SiN-caps has a similar aging rate in air to that of a no-cap device in hydrogen ambient. This is due to the hydrogen associated with the SiN-caps [2]. When a device with SiN-caps is radiated and heated in air we observe a similar build-up of fast states to that shown by curve B in Figure 3 [6].

In hot-hole degradation, the fast states first appear at high-current levels and, as the aging continues, they move to lower and lower energies [5]. The same process is also observed in the radiation damage mechanism, which is illustrated in Figure 6. A large area MOSFET with SiN-caps is first exposed to radiation and then heated in air at 100 ° C. The curve 1 is the post-irradiation drift which shows uniformly trapped holes and few fast states. The curves 2 through 5 show a gradual reduction of holes and a monotonic increase in the fast states as a result of heating. Note that the fast states appear at the high current levels first and then move toward the lower and lower current levels, in a manner similar to that in the hot-hole injection mechanism [5].

The radiation damage has long been known to be self-limiting [7]. Recently, we have shown that the hot-hole damage is also self-limiting [5]. The extent of damage to the device depends on τ_c which in turn depends on bias conditions, ambient gases and temperature, and the processing conditions. With the growing number of implant and dry processing steps and simultaneously reduced processing temperatures, the density of neutral traps in the interface region has steadily increased [8, 9]. These traps, we believe [9], are responsible for the observed increase in the rate of build-up of fast states in both the mechanisms.

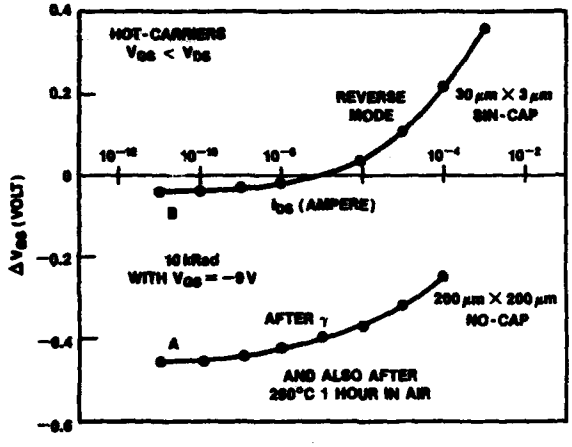


FIGURE 4: *Comparison of hot-carrier drift ($V_{GS} < V_{DS}$) in reverse mode with radiation induced drift ($V_{GS} = -9V$).*

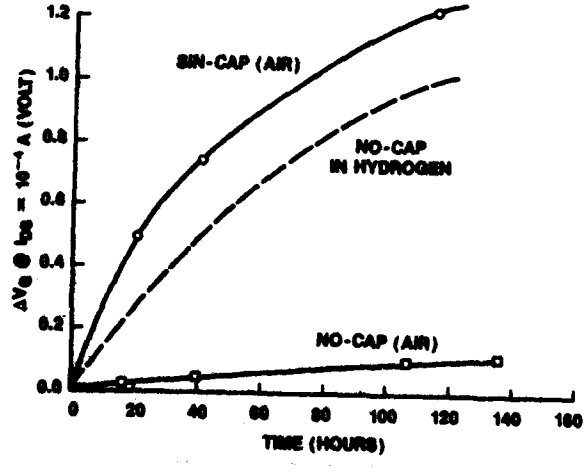


FIGURE 5: *Role of SIN-Caps in hot-carrier aging.*

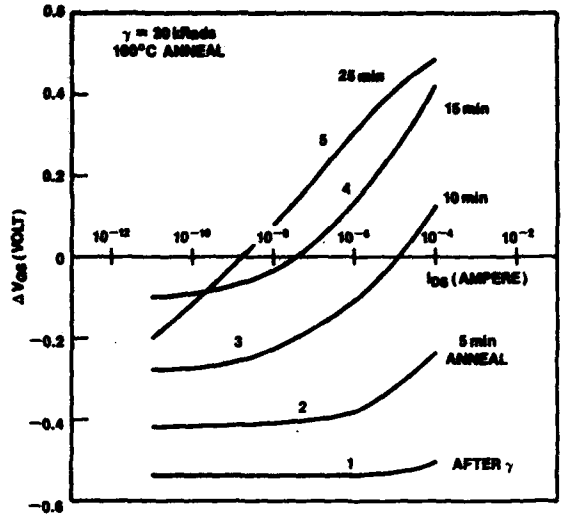


FIGURE 6: *Conversion of holes to fast states in 200μm x 200μm MOSFET with SiN-Caps.*

The radiation damage to DRAMs manifests in ways similar to the hot-carrier damage. The main difference in the two mechanisms is that the radiation affects all the devices on a DRAM chip whereas the hot-hole injection degrades only a few selected devices. The V_{DDmin} degradation due to radiation, which is shown in Figure 7, is similar to that observed due to hot-carrier injection [10]. The

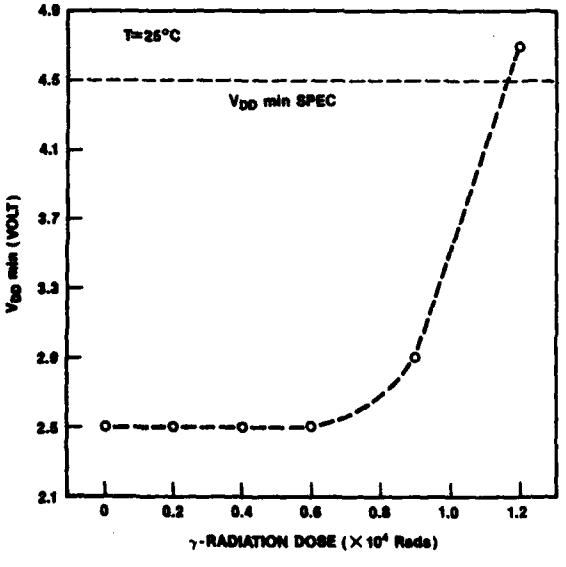


FIGURE 7: *A typical degradation of V_{DDmin} of DRAMS operating under bias during irradiation.*

V_{DDmin} is the minimum supply voltage required by a DRAM to operate under a given set of timing conditions.

A direct quantitative mathematical correlation between the two mechanisms is rather difficult, mainly because of the two dimensionality associated with the hot-hole degradation of the short-channel devices. Nevertheless, qualitatively the two mechanisms have similar effect on device characteristics. The kinetics of hole trapping and their conversion are identical in both the cases. Therefore the response of the interface to hot holes can be conveniently explored by radiation damage experiments.

4. CONCLUSIONS

The radiation and hot-hole damages are related to each other through the kinetics of hole trapping and conversion to fast-states at the interface. Qualitatively, the presence of hydrogen in the ambient and the processing conditions have similar impact on the responses of the devices to these two degradation mechanisms. For a given probability of hot-hole generation and their injection, a device with increased radiation softness would have a higher susceptibility to hot-carrier damage.

REFERENCES

[1] T. H. Ning, Solid-State Electronics, Vol. 21, pp. 273-282, 1978.

[2] R. C. Sun, J. T. Clemens, and J. T. Nelson, Proc. 18th International Reliability Phys. Symp., Las Vegas, Nevads, pp. 244-251, April 1980.

[3] K. K. Ng and G. W. Taylor,39th Annual IEEE Device Research Conference, Santa Barbara, CA, June 22-24, 1981, p. IVB-1.

[4] A. G. Sabnis and J. T. Nelson, Proc. 21st Interntional Reliability Phys. Symp., Phoenix, Arizona, pp. 90-95, April 5-7, 1983.

[5] A. G. Sabnis and J. T. Nelson, IEEE IEDM Tech. Digest, pp. 52-55, Dec. 1985.

[6] A. G. Sabnis, IEEE Transactions on Nucl. Sci., Vol. NS-30, No. 6, pp. 4094-4099, Dec. 1983.

[7] J. P. Mitchell, IEEE Transactions on Electron Devices, ED-14, 764 (1967).

[8] R. A. Gdula, IEEE Transactions on Electron Devices, Vol. ED-26, No. 4, pp. 644-647, April 1979.

[9] A. G. Sabnis, Proc. 22nd International Reliability Physics Symposium, Las Vegas, Nevada, pp. 156-160, April 3-5, 1984.

[10] A. G. Sabnis, Proc. Int. Conf. on Semiconductor and IC Technology, Beijing, China, Oct. 20-23, 1986, pp. 660-663.

The role of holes and electrons in the aging of MOS transistors

M. Tosi L. Baldi F. Maggioni
SGS Microelettronica
via C.Olivetti 2
20041 Agrate Brianza (MI)-Italy

A new kind of representation of aging effects in MOS transistors on a V_D/V_G diagrams is proposed. The working curve of N-channel or P-channel transistors, for different loading factors, can be easily plotted in this diagram and the regions of hole or electrons injection during the normal operating cycle identified. Since any change in transistor technology is reflected in a change of the iso-aging diagram, this kind of representation allows to compare different technologies and their limits with the need of different circuit configuration. The results of experimental evaluation of different LDD structures are discussed.

1 Introduction

Electron injection in N-channel and P-channel de[illegible] has been widely studied as main cause of the degradation in transistor performances. Also hole injection has been recently demonstrated under particular bias conditions [1], [2].

The most evident effect of carrier injection is the variation in transconductance and/or threshold voltage of the stressed device according to the technology and the stress conditions. Both increase or decrease of these critical parameters can be obtained, therefore different authors have reported different aging behaviors, and there is no general agreement on which the worst case stress conditions are. Moreover, even if a very exhaustive literature exists on aging effects, it is very difficult to get a clear and coherent picture and to compare results obtained by different authors.

We propose a new kind of representation of aging effects on a V_D/V_G diagram, taking into account all possible device bias conditions. In this way, the aging characteristics of different technologies can be easily compared on the same diagram and superimposed to the working curve of critical devices, as obtained by SPICE simulations, showing which are the dominant injection mechanism in the different regions.

2 Experimental

To obtain these diagrams, the transistor transfer characteristic in the linear region (at $|V_D| = 100mV$) has been measured with an HP4145 Parameter Analyzer before and after 10 minutes stress. An example of this curves is given in Fig.1 for a N-channel device. An increase in transconductance (Fig.1a) is the consequence of majority (with regard to the bulk) carrier trapping, while a decrease (Fig.1b) means trapping of minority carriers. Symmetrical results are obtained for P-channel transistors. As limit of the no-injection area, an arbitrary limit of less than 1% variation in transconductance over a 10 minutes stress has been assumed. Of course tighter limits could be imposed, by increasing the stress time. Also iso-aging curves can be drawn to evidence the regions where the effects are more marked.

An example of this kind of representation is given in Fig. 2 for the N-channel and P-channel transistors of a 1.2 μm CMOS N-well. The gate oxide thickness of the devices is 28 nm, the threshold voltage around +0.85V and -0.85V respectively and an LDD structure has been used for the N-channel alone.

As discussed in the introduction, both hole and electron trapping can be obtained with different bias con-

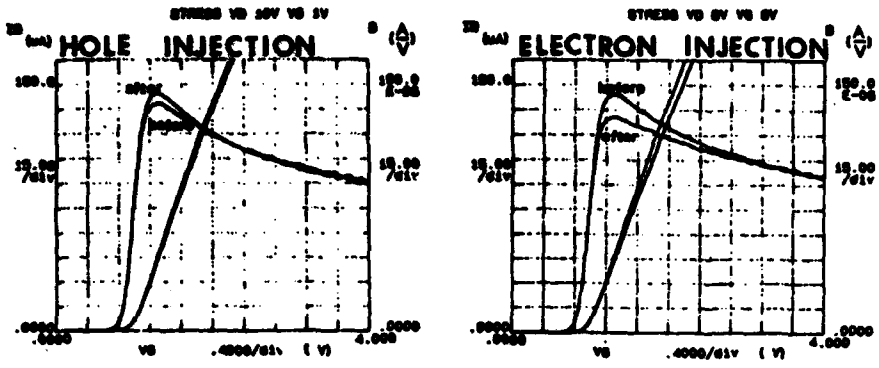


Fig. 1

ditions. For P-channel devices, electron trapping takes place at low gate voltage, while hole injection is evident at high gate voltage. For N-channel devices, there is hole trapping only for high drain and low gate voltage, while electron trapping takes place at high gate voltages. It is rather interesting that electron trapping is evident at low drain and low gate, where the electric field configuration tends to oppose electron injection into the gate. This could be explained by electron trapping into interfaces states, rather than into the bulk of the gate oxide. Since however the effects on the electrical parameters are the same, this region will be included in the electron injection region.

3 Comparison between different process architectures

This kind of diagram can be easily used to understand how the aging characteristics of a process architecture may influence the operation of devices.

The ring oscillator, sketched in Fig.3, has been simulated with SPICE for a 10V bias voltage, and the operating curves of the N-channel and P-channel transistors have been superimposed as dashed lines on the aging diagrams in Fig.2. From this figure it can be noted that, in this case, only electron trapping is important for both transistors. No aging takes place during the switch on of either transistor (the lower branch of the device working curve), while during the switch off (the higher branch), electrons are injected both in N-channel and in P-channel transistors.

Inside each injection zone different degradation mechanisms can be active, depending on the bias conditions and the devices architecture. As has been shown by simulation and experimental results [3] the carriers can be injected over the drain-gate overlap region or over the channel, producing different variations in the electrical parameters. For example, while carriers injection over the channel induces a shift in threshold voltage, charge trapping over the drain is associated only with transconductance variations.

The degradation mechanism is strongly dependent also on the process architecture. LDD and not LDD devices behave in quite different ways, and even among LDD structures, strong differences in aging can be evidenced with different n- concentrations.

As an example two different N-channel LDD have been measured, with high and low n- doping concentration, corresponding to an n- implant dose of $1 \cdot 10^{14}$ ions/cm^2 and $1 \cdot 10^{13}$ ions/cm^2 respectively. To simulate devices operation, six points lying over the switch off curve of the N-channel device of the ring oscillator have been considered (the points are marked with circles in Fig.2a). Transfer characteristics have been measured before and after a two minute stress in each point. Most of the aging takes place in the last two stress points, at $V_D = 10$V and $V_G = 1{,}2$V. As shown in Fig.4 the effect of the stress is electron trapping over the channel for the high n- concentration LDD, and electron trapping over the drain for the low n- concentration LDD. The different aging mechanism can be easily explained by the different configurations of the electric field in the channel, due to the different concentration profiles as shown in Fig.5. MINIMOS simulations show that for high n- doping concentration only one peak of electric field is present, at the

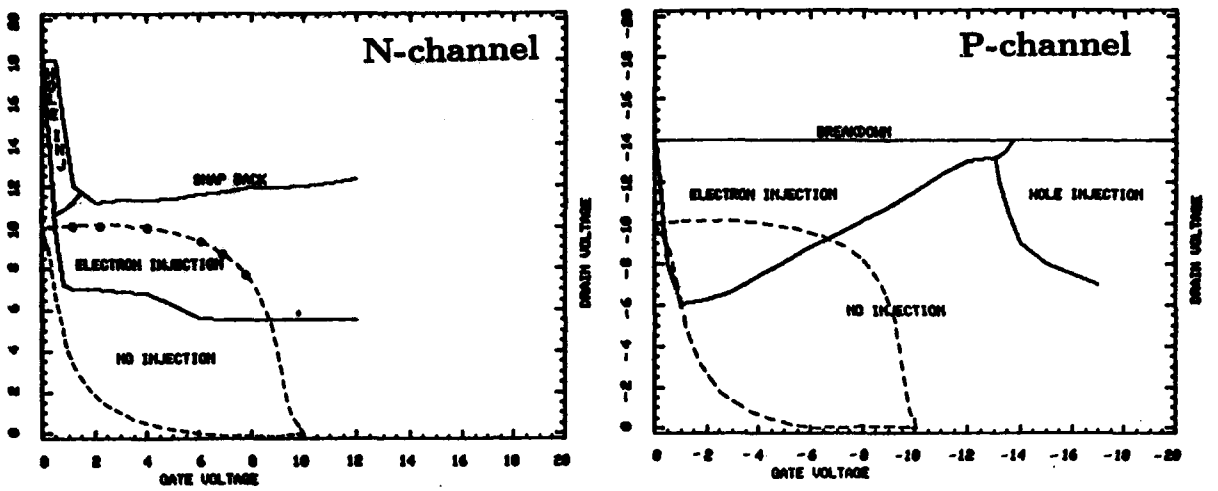


Fig. 2 Iso-aging diagrams and ring oscillator SPICE simulation.

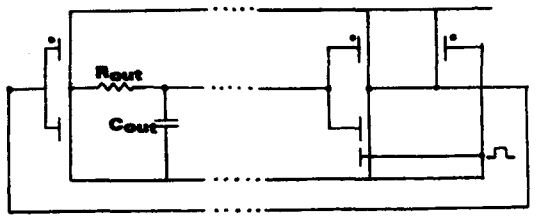


Fig. 3 CMOS, seven stages ring oscillator scheme.

channel/n- interface, while for low n- doping concentration, two peaks are present, the highest at n+/n- interface. In the first case, electron are injected over the channel, altering the threshold voltage, while in the second case, electrons are injected over the drain, changing only the transconductance. If the gate voltage is increased, the electric field at n+/n- junction decreases, as shown in Fig.6, and therefore the same aging effects are evident in the two cases. The knowledge of aging mechanism in every diagram point is very important specially for those devices which work in both hole and electron injection region and where a partial recovery is possible. This is the case for example of a V_{pp}/V_{cc} switch, used in EPROM Memories, whose drain voltage varies, at low gate voltage, between supply voltage and programming voltage.

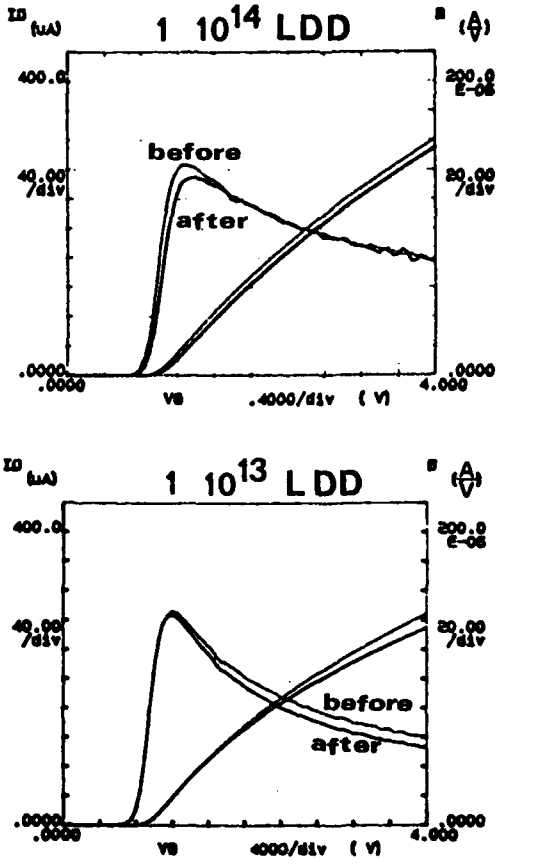


Fig. 4 Comparison between the behaviour of the two kinds of LDD transistors after the ring oscillator switch off stress.

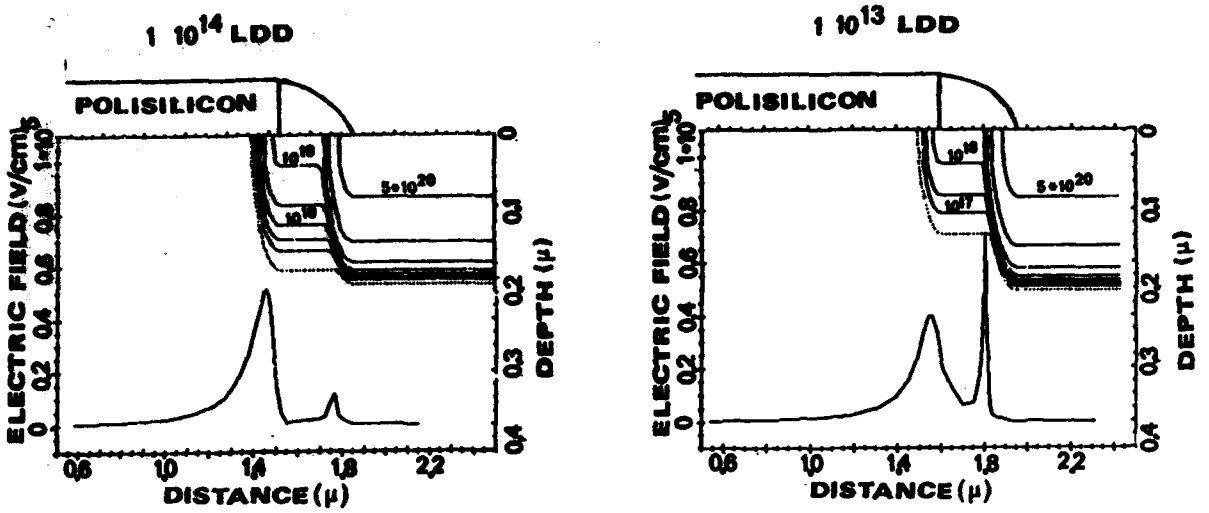


Fig. 5 MINIMOS simulation of the electric field at the stress bias condition and MINIMOS simulation of drain isoconcentration lines.

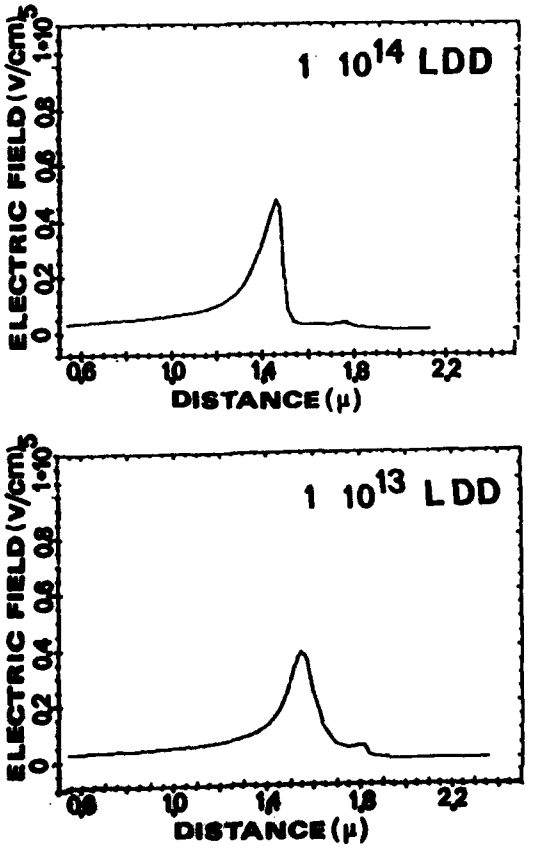


Fig. 6 MINIMOS simulation of the electric field at high V_D and high V_G

4 Conclusions

The MOS transistors degradation takes place as consequence of both holes and electrons injection in gate oxide. A correct way to represent aging effects is to plot them on a V_D/V_G plane over which also the devices working curve can be plotted and both the global and the point by point devices behaviour in the normal operating cycle can be studied.

Both the value and the mechanism of degradation are strongly dependent on the technology; the use of this kind of diagrams allows the comparison between different device architectures over all the operating range of critical device.

Acknowledgments

This work has been supported by ESPRIT project 554-SPECTRE. We should like to thank Ing. Roberto Mucciante for the aid in SPICE device simulation.

References

[1] Tzou et al. IEEE Elec. Dev. Letters EDL-7, 5 (1986)

[2] Hofmann et al. IEEE Trans. Electron Dev. ED-32, 691 (1985)

[3] Bergonzoni and Doyle "Simulation of aging effects in MOS transistors" ESSDERC 1987

THE VOLTAGE DEPENDENCE OF DEGRADATION IN N-MOS TRANSISTORS

B.S.Doyle, M.Bourcerie, J-C. Marchetaux & A.Boudou

BULL S.A., Ave. Jean-Jaures, 78340 Les Clayes Sous Bois, FRANCE

Hot carrier stressing has been carried out on silicon n-MOS devices as a function of gate voltage, at fixed drain voltages. It is found that a maximum of degradation occurs not only at Vg=Vd/2, but also at Vg=Vd. It is further found that the time power law for threshold voltage shift changes according to the voltage ratio. It is suggested that while the first peak is due to interface state degradation, the second is due to another process, possibly electron trapping in the oxide.

1. INTRODUCTION

In the study of hot carrier stressing in silicon n-MOS transistors, it has long been recognised that the effect of stressing at high voltages results in a degradation in time of the transistor I-V characteristics. This degradation is usually characterized by measuring the decrease in transconductance (g_m) or the increase in threshold voltage (V_t). These degradations follow a power law as a function of time, and it is thus possible to extrapolate to a certain level of Vt shift or gm degradation and define a lifetime for the device. Carrying this out at different drain voltages, it becomes possible to extrapolate back to power supply voltages, and estimate the lifetimes in the circuit working environment.

In order to do this, it is necessary first first to study the gate voltage dependence of aging in order to choose the maximal conditions of aging. It has been found that maximum aging occurs in the saturated region of the transistor characteristics, at Vg=Vd/2. This corresponds also to the maximum of the substrate current. The coincidence of the maximum of the substrate current and the maximum degradation has led to models in which substrate current alone [1] or substrate current combined with the drain current [2] are sufficient in order to quantify the susceptibility of any given transistor to stress degradation. In this presentation, it will be shown that the situation is more complex. It will be shown that maximum degradation does not always occur at Vg=Vd/2, suggesting that there is another type of damage occuring in the device during hot carrier stressing.

2. EXPERIMENTAL

Measurements have been carried out principally on the 2μ devices, with width 50u. The transistors were fabricated on 15 ohm.cm substrates with a threshold implant of $1.4*10^{12}/cm^2$. The gate oxide had a thickness of 40nm, was grown at 1000°C, followed by a Post Oxidation anneal at 1000°C for 30 minutes. The Post Meatlization Anneal was carried out in a nitrogen-hydrogen atmosphere, with a passivation layer of SiO_2.

The measurements were performed on a microcomputer-controlled HP 4145 Semiconductor Parameter Analyser. The threshold voltage (V_t) in each case here is defined as the gate voltage necessary for a drain current of 1μA/μ gate width.

3. RESULTS AND DISCUSSION

Figure 1 shows a typical plot of the threshold voltage shift as a function of gate voltage at which the degradation was performed. Also included is the substrate current Vs gate voltage for the same drain voltage as the aging. It can be seen that the maximum occurs at the same point as the $I_{sub(max)}$, as has been discussed earlier.

This, however, is not always the case.

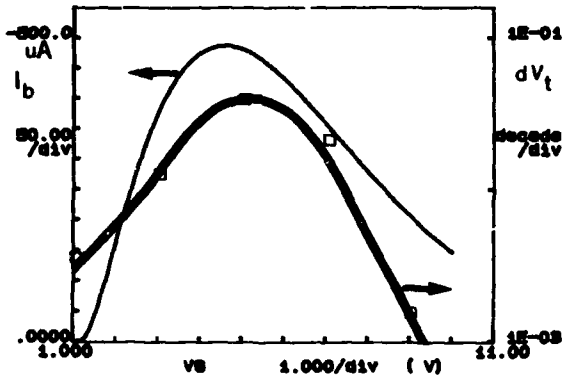


FIGURE 1

Substrate current and threshold voltage shift versus gate voltage at Vd=9.5V for a 2μ device, showing a maximum in both cases at Vg=Vd/2.

Figure 2 shows the threshold voltage shift after 1000 seconds for conventional 2 μ devices in the full gate voltage range. It can immediately be seen that two peaks appear in the plot. The first peak occurs at Vg=Vd/2 and is identical to that of figure 1. The second peak occurs at Vg=Vd, at gate and drain voltages not normally associated with aging. These voltages correspond to MOS functionning in the linear regime, with an inversion layer extending from the source to the drain, and low transverse fields. That the transverse field is low can be seen from figure 1, where the substrate current at Vg=Vd is one fifth that at Vg=Vd/2.

Looking at the gate currents, represented in figure 2, it can be seen that the the maximum of the first δVt peak corresponds to the point at which gate hole and electron currents are equal. This observation and other measurements [3] has lead to a model in which the creation of interface states results from the presence of both holes and electrons together in the oxide.

In the case of the second δVt peak however, it can be seen that the hole gate current is negligible and the electron gate current is maximum at this point, indicating that the hot carrier damage is in some way related solely to electrons. This suggests that the degradation is due to electron trapping in the oxide, as the conditions resemble those of capacitor photo I-V measurements (small fields in the oxide between drain and gate and high electron concentrations) and this is known to result, at high electron fluxes, in electron trapping in the oxide [4].

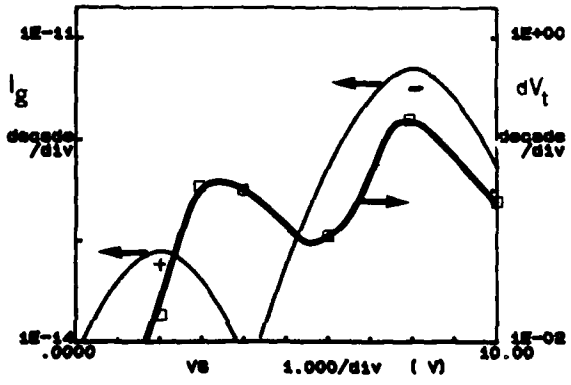


FIGURE 2

Threshold voltage shift versus gate voltage for a 2μ conventional device, aged at Vd=8V showing a double maximum. A representation of the gate current at the same drain voltage (+ for holes, - for electrons) is also shown.

It is thus proposed that the double δVt peak results from different processes, interface state generation and charge trapping. If these are indeed different processes it might be expected that there would be some difference in the evolution of the degradation with time. Figure 3 shows the dVt Vs time shifts at the voltage conditions corresponding to the two peaks. It can be seen that they both follow a power law, of form

$$dVt=At^n,$$

each voltage ratio having its own 'A' and 'n'. It can further be seen

that the greater gradient is associated with the interface state creation and has values of about 0.65, whereas the degradation at Vg=Vd has the lesser slope, around 0.25. This net difference in slopes shows that the charging kinetics in each case is different, and is a further indication that the damage created at these two voltage ratios is different.

Although the results of figure 3 are not seen in on all fabrication line devices, they are seen more widely than might be thought.

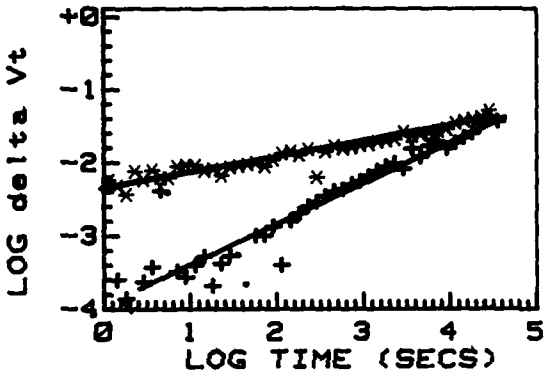


FIGURE 3

Threshold voltage shift with time plot for a 2μ conventional transistor, aged at Vg=Vd/2 (+ symbols) and Vg=Vd (* symbols), showing the presence of quite seperate power laws slopes, depending on the voltage ratio.

We have seen this effect on devices from five out of six pilot lines whose stress integrity we have tested. Figure 4 gives an example of the double power law seen from another pilot line. In this particular case, the devices have been fabricated using electron beam technology. It can be seen that once again the Vg=Vd/2 stressing has a power law gradient about 0.6, while the devices stressed at Vg=Vd has a power law of the order of 0.3.

Given that the damage at different voltage ratios results from different mechanisms, it would be expected that the pre-time power law factor A and the gradient 'n' be different

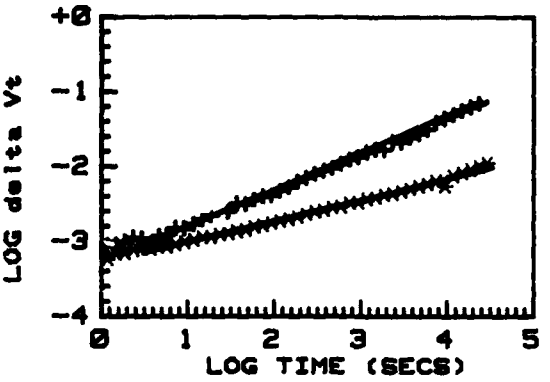


FIGURE 4

Threshold voltage shift with time plot for a different pilot line device, aged at Vg=Vd/2 (+ symbols) and Vg=Vd (* symbols), showing once again, the presence of quite seperate power law slopes, depending on the voltage ratio.

for each technology, and depend on the process steps. If the damage at Vg=Vd results from trapping of electrons on defects already present in the oxide, it could be imagined that under certain circumstances, a combination of both these power laws would be seen, depending on the gate-to-drain voltage ratio and the process conditions. Figure 5 shows the threshold voltage shift as a function of time for a third pilot line device. The voltage conditions here are Vg=Vg/2. It can be seen that there are two clear time law depencencies, the first having a slope 0.3 and extending over 3 1/2 decades in time, and the other having gradient 0.8 and dominating the long stressing times. It is obvious that the summation of two time power laws results in one power whose slope depends on which of the power laws dominates at the stressing time in question. The fact that the smaller gradient of 0.3 which could be associated with electron trapping in the oxide is seen initially, at Vg=Vd/2, could be interpreted simply as meaning that the pre-time power law factor 'A' for this behaviour is great, indicating that this

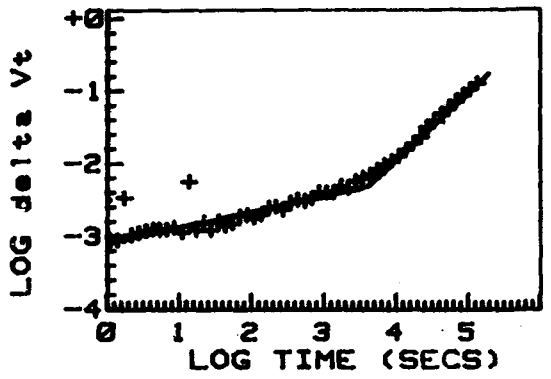


FIGURE 5

Threshold voltage shift with time plot for a LDD 1.3u transistor,stressed at Vg=Vd/2 and showing the presence of two power laws in the same time plot.

particular oxide is more sensitive to electron trapping than others.

4. CONCLUSIONS

Hot carrier degradation has been carried out on 2u n-MOS devices, at constant drain voltage, with gate voltage as the variable. It has been found that the degradation peaks at two points, Vg=Vd/2, and Vg=Vd. The fact that the second peak occurs at the point of maximum gate electron current and negligible gate hole current, indicates that the damage in this case is different to the interface state generation damage associated with the Vg=Vd/2 maximum and might possibly be due to electron trapping.

REFERENCES

[1] Takeda, E. and Suzuki, N., IEEE Electron Device Lett. EDL-4 (1983) 111.

[2] Weber, W., Warner, C. and Scherin, A.V., IEDM Proceedings (1986) 390.

[3] Saks, N. et al. IEEE Trans. Electron Dev. ED-33 (1986) 1529.

[4] Nicollian, E. and Brews, J.R., The Physics of MOS Devices (J. Wiley and Sons, New York, 1984).

Characterization and analysis of hot-carrier degradation in p-channel transistors using constant current stress experiments.

R. Bellens, P. Heremans, G. Groeseneken, H.E. Maes

IMEC, Kapeldreef 75, B-3030 Heverlee, Belgium

An alternative procedure is proposed to perform accelerated lifetime experiments in p-channel transistors in order to resolve some problems that occur using the conventional constant voltage procedure. The results are correlated with the degradation mechanisms and a comparison with n-channel transistors is carried out.

INTRODUCTION

In order to predict the lifetime of MOS-transistors under influence of hot-carrier degradation and to intercompare the hot-carrier sensitivity of different MOS technologies, the linear relationship between $\log[I_{sub}/I_d]$ and $\log[\tau*I_d]$ or $\log[\tau*I_{sub}]$ for n-channel transistors and p-channel transistors, respectively, (defined as the lifetime curve) has commonly been applied [1,2,3]. In these experiments, the lifetime τ is related to an arbitrary chosen shift of the threshold voltage [1] or transconductance [2] during a constant voltage stress condition.

In this work, accelerated lifetime experiments (ALE) were performed on both n- and p-channel transistors based on these relationships. It is shown that for p-channel transistors, this procedure can however not be applied as such due to the influence of trapped charge on I_{sub} and I_d. In order to solve this problem an alternative ALE-procedure, using constant current stress experiments, is proposed.

The threshold voltage shifts (ΔV_t) and the corresponding variations of I_{sub} and I_d with time for several stress conditions are analysed and discussed for both transistor types. These changes are correlated with the different degradation mechanisms that occur depending on stress condition and channel type. This leads to important conclusions on the interpretation and the validity range of the ALE-procedure for both n-channel and p-channel transistors.

EXPERIMENTS and DEVICES

For this study MOS transistors from different suppliers were used, with channel lengths ranging from 1.75 to 3µm, oxide thicknesses from 27 to 40nm and channel width of 25µm.

The threshold voltage for all experiments and devices was defined as the gate voltage at which drain current is 0.2µA per µm width at V_{ds}=5V. The threshold voltage was measured in reverse operation, i.e. by interchanging source and drain after stressing. The device lifetime τ was defined as the stress time necessary to obtain 10mV threshold voltage shift.

I. A.L.E. IN N-CHANNEL TRANSISTORS

Several n-channel transistors were stressed in the region of maximum degradation, i.e. near the maximum of the substrate current. In this region the threshold voltage increases during stress with a slope of typically 0.7, without showing any considerable saturation tendency, as shown on fig. 1. The lifetime τ can be described by a linear relationship between $\log[\tau*I_d]$ and $\log[I_{sub}/I_d]$, with a slope of typically 2.7, which corresponds with the values found in the literature [1,3]. This is shown on fig. 2 which compares the hot-carrier sensitivity of the different technologies used in this study.

II. A.L.E. IN P-CHANNEL TRANSISTORS

For the case of p-channel transistors, the behaviour is quite different. As shown on fig. 1, the slope of ΔV_t vs. time is, for a p-channel

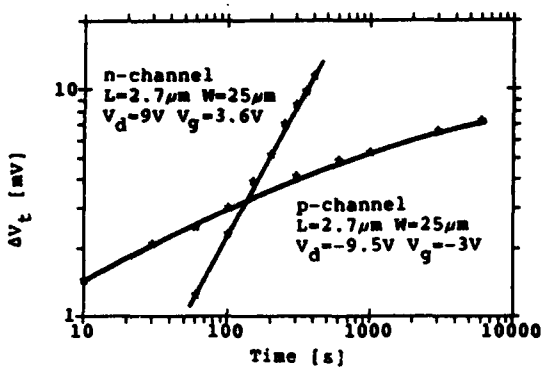


Fig. 1: ΔV_t vs. time for p and n-channel transistors

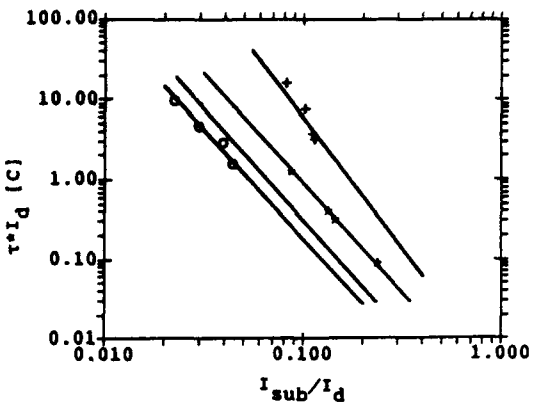


Fig. 2: Intercomparison of hot-carrier sensitivity of different technologies for n-channel transistors

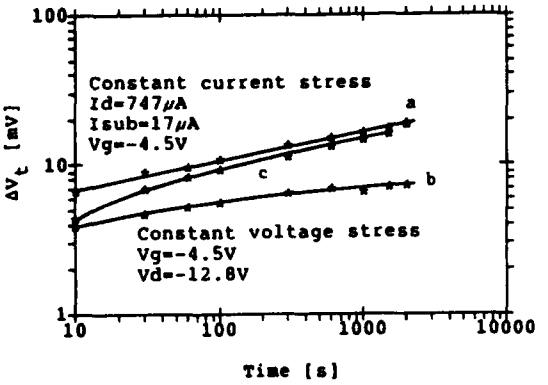


Fig. 3: Comparison between constant current, constant voltage and gradually adjusted voltage stress on p-channel transistors

transistor stressed under condition of maximum substrate current, only 0.2. Moreover, ΔV_t tends to saturate. As a result of this saturation, linear extrapolation of the curves and from it an extraction of the lifetime proves to be hazardous and totally unreliable. It indeed becomes difficult to obtain the $\tau * I_{sub}$ vs. I_{sub}/I_d relationship that is expected from the simple hot-carrier model [1] and intercomparison of the hot-carrier sensitivity of different technologies is almost impossible.

This phenomenon is explained by the degradation mechanism of p-channel transistors. Indeed, in p-channel devices the degradation for stress at the maximum substrate current is known to be predominantly due to electron trapping in the oxide near the drain [4,5]. This causes a local reduction of the lateral electric field and consequently a considerable decrease of the substrate and gate currents during stress. These effects evidently result in a reduced electron trapping and a corresponding saturation of ΔV_t.

III. A.L.E. USING CONSTANT CURRENT STRESS

Based upon this understanding, an alternative stress procedure is proposed and used: the p-channel transistors are stressed by imposing a constant current at the drain and at the substrate while keeping source and gate voltage constant. This ensures a constant electric field near the drain, which should eliminate the saturation effect.

Comparing the results of constant current (fig. 3, curve a) and constant voltage stress (fig. 3, curve b), one can observe two main features. Firstly the slope of the ΔV_t vs. time is about 0.2 in both cases and secondly the saturation effect observed for curve b has disappeared for the case of constant current stress, as expected. Notice that curve b lies below curve a, which is due to the larger electron trapping during the constant current stress, as the result of the continuous increase of the applied voltages (V_d and V_{sub}). In order to check the validity of this conclusion a third experiment was performed. Constant voltage stresses were applied but the stress voltages (V_d and V_{sub}) were adjusted at each measurement time to the values

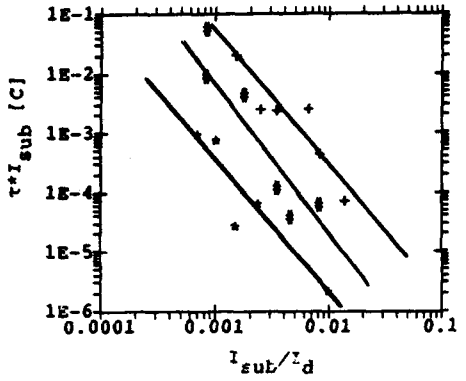


Fig. 4: Intercomparison of hot-carrier sensitivity of different technologies for p-channel transistors

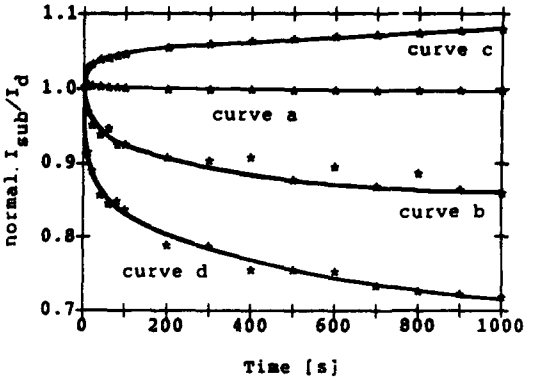


Fig. 5: Isub/Id, normalised to the initial value, during constant voltage stress for n-channel (Vd=8.5V curve a: Vg=3V, curve b: Vg=.8V, curve c: Vg=Vd) and p-channel transistors (curve d: Vg=-3V Vd=-8.5V) L=1.75μm W=25μm

measured during the constant current stress at the corresponding time. The result of this experiment is represented by curve c on fig. 3. As can be seen, ΔV_t is at first identical to the value for the constant voltage case, while the curve gradually approaches the one for constant current stress.

P-channel transistors from different technologies were evaluated using this new stressing procedure. Now, the threshold voltage shifts result in the expected $\tau * I_{sub}$ vs. I_{sub}/I_d relationship with a slope of about 2.3, as shown on fig. 4. In this way it becomes possible also for p-channel devices to reliably intercompare the hot-carrier sensitivity of different technologies (fig. 4). But of course the procedure does not yield a useful lifetime prediction for constant voltage conditions.

IV. COMPARISON with N-CHANNEL TRANSISTORS

As demonstrated in the previous section, charge trapping during stress can strongly disturb the conventional ALE-procedure because it changes the I_{sub}/I_d ratio, which is the abscissa of the lifetime curve. The question arises why this phenomenon was never observed or reported for n-channel devices. Therefore a thorough analysis of the degradation of n-channel transistors at different stress conditions has been carried out, and a comparison with p-channel devices was made. This is illustrated on fig. 5 and fig. 6.

On fig. 5 the ratio I_{sub}/I_d during stress is plotted as a function of the stress time for both n- and p-channel transistors, stressed under different constant voltage conditions which are indicated on the figure. On fig. 6, ΔV_t vs. time is plotted for both channel types under different conditions, for voltage as well as for current stress.

The case of the p-channel transistors is illustrated by curve d on fig. 5 and 6. I_{sub}/I_d decreases significantly, due to build-up of a net negative trapped charge, as already concluded from the previous section.

However, as can be seen on fig. 5, for the case of a n-channel device stressed around the maximum of the substrate current, I_{sub}/I_d remains constant during stress (curve a), although both I_{sub} and I_d decrease in time. This confirms the results found in other studies [5], in which the dominant degradation mechanism for this case was shown to be interface state generation. Indeed, this causes a decrease of I_d, due to mobility degradation, and consequently a correlated decrease in I_{sub}. However, since no considerable trapping occurs, the multiplication factor (I_{sub}/I_d) remains constant. As seen on fig. 6, no saturation in ΔV_t vs. time is observed and constant current and constant voltage stress yields identical results (curve a and e). The

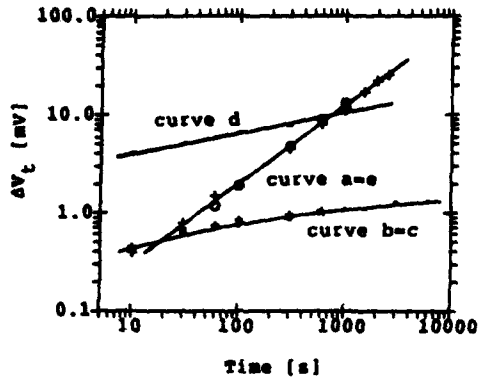


Fig. 6: ΔVt vs. time for n-channel (curve a: Vg=4.2V Vd=9V, curve b: Vg=.8V Vd=9V, curve c: Vd=Vg=7.5V, curve e: constant current stress) and p-channel transistors (curve d: constant current stress) L=2μm W=25μm

slope of 0.7 found for this stressing condition (max. I_{sub}) seems to be typical for this degradation mechanism.

For the case where $V_g \approx V_{th}$ (V_g=0.8V), the multiplication factor decreases (fig. 5, curve b), which means that a net positive charge is trapped. This is again in agreement with [4,5], where it was shown that the dominant degradation mechanism in this case is hole trapping, leading to a negative ΔV_t. For this case a saturation effect of $|\Delta V_t|$, correlated with the hole trapping is indeed observed (fig. 6, curve b). The slope of 0.2, similar to the p-channel case, seems to be typical for the trapping mechanism.

Finally, stressing the n-channel transistors at $V_g \approx V_d$ results in an increase of the multiplication factor (fig. 5, curve c), which is indicative for a net negative charge trapping. Furthermore the slope of the ΔV_t vs. time curve is also 0.2 and a saturation is again observed (fig. 6 curve c). These phenomena point to an electron trapping mechanism as the main degradation cause for $V_g \approx V_d$.

For the conditions, depicted in curve b and c of fig. 5 and 6, the conventional ALE-analysis is subjected to the problems discussed in section II and can be replaced by the alternative constant current procedure.

V. CONCLUSIONS

The evaluation of the relationship between I_{sub}/I_d and $\tau * I_{sub}$ for p-channel transistors under I_{submax} stress conditions is shown to be strongly disturbed by saturation phenomena in ΔV_t vs. time and a corresponding decrease of I_{sub}/I_d. These effects are due to electron trapping and can be circumvented by performing constant current stress experiments.

Similar effects occur in n-channel transistors, when trapping is the dominant degradation mechanism i.e. for $V_g \approx V_{th}$ or $V_g \approx V_d$. These problems do not occur for those conditions where the degradation is dominated by the creation of interface states ($V_g \approx V_d/2$).

Therefore, care should be taken when using the relationship between I_{sub}/I_d and $\tau * I_{sub}$ or $\tau * I_d$ for p-channel and n-channel devices, respectively, as a tool to predict the device lifetime under constant voltage conditions. The ALE-evaluation can however be applied for intercomparison of the hot-carrier sensitivity of different technologies, provided the appropriate selection of either constant voltage or constant current stress is made.

ACKNOWLEDGEMENT

The authors wish to thank BTMC Alcatel for providing the devices.

P. Heremans is a research assistant and G. Groeseneken a senior research assistant of the Belgian National Fund for Scientific Research.

REFERENCES

[1] C. Hu et al., IEEE Trans. on Electron Devices, vol. EDL-32, No. 2, pp. 375-385.
[2] J.J. Tzou et al., IEEE Electron Device Letters, vol. EDL-6,No. 9, pp. 450-452.
[3] W. Weber et al., IEDM Techn. Dig.,pp. 15.4, 1986.
[4] P. Heremans et al., Coll. IEE on "Hot Carrier Degradation in Short Channel MOS" London, jan. 87.
[5] P. Heremans et al., SISC 1986
[6] P. Heremans et al., IEEE Electron Device Letters, vol. EDL-7, No. 7, pp. 428-430.

CORRELATION BETWEEN FLATBAND VOLTAGE SHIFT IN MOS CAPACITORS AND ENDURANCE DEGRADATION OF EEPROM CELLS *

Jacek MANTHEY, Michel DUTOIT, Marc ILEGEMS

Institute for Microelectronics, Federal Institute of Technology, 1015 Lausanne, Switzerland

Charge trapping in thin injection oxides used in EEPROMs is studied as a function of charge injected into the oxide. In MOS capacitors, pulsed alternating current injection is used to simulate the operation of EEPROM. The resulting shifts of flatband voltage are correlated with the shifts of threshold voltage in the WRITTEN and ERASED states in EEPROM cells. Positive charge generation depends more than negative trapping on the negative field strength and is found to vary with the stress history of the oxide.

1. INTRODUCTION

The degradation of thin oxides used in floating-gate nonvolatile memories (EEPROM) has often been studied by constant current injection in MOS capacitors. However, in real operation, these oxides are stressed by pulses of alternate polarity [1]. The purpose of this work is to compare charge trapping during DC and pulsed AC current injection. The relative roles of positive and negative pulses, their amplitudes and durations are shown. There is good correlation between these measurements and the wearout of EEPROM cells.

2. EXPERIMENTAL

We studied silicon gate MOS capacitors (area 10^5 um^2) and p-channel EEPROM cells (thin oxide area 9 um^2) fabricated with a production 3um p-well CMOS process [2]. Thin oxides were grown to 12 nm on p-doped substrates (N_A = 2x10^{18} B/cm^2) at 950°C in N_2-diluted dry O_2 with 3% HCl. The capacitors were stressed with constant current pulses using a Keithley 220 programmable current source. The amplitudes and durations of the positive and negative pulses were varied independently to study the influence of field polarity and strength on charge trapping. Periodically, stressing was interrupted in order to record high-frequency C-V curves with a HP4275 LCR meter.

* This research was supported by the Swiss National Foundation for Scientific Research (Program No.13) and the Commission for the Encouragement of Scientific Research.

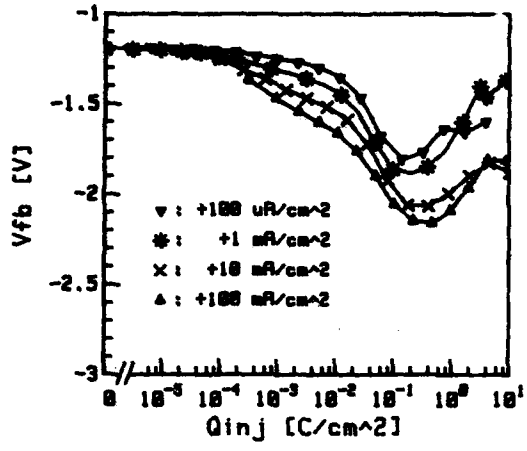
For EEPROM cells, the threshold voltages V_{th} of the written and erased (W/E) states and thus the memory window (the gap between V_{thW} and V_{thE}) were monitored as a function of the number of W/E cycles. The peak electric field was varied by adjusting the W/E pulse risetimes [3]. These measurements were performed with two HP8112 pulse generators and a HP4145 semiconductor parameter analyzer.

Both experiments used a HP3488 switch unit and were controlled by a HP9826 computer .

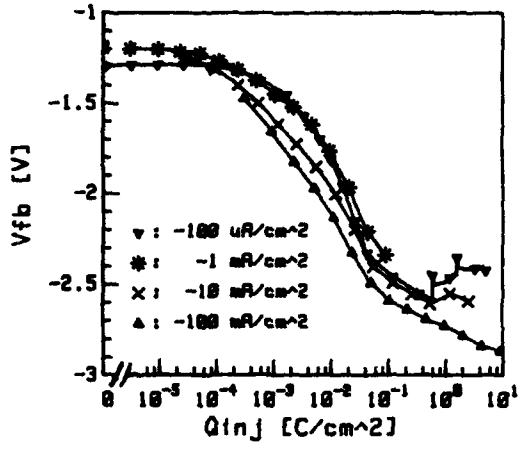
3. RESULTS AND DISCUSSION

The variation of the flatband voltage V_{FB} during injection is shown in figs.1-2. For a DC stress, positive charges are generated at low Q_{inj} (initial negative shift of V_{FB}), as was previously reported [4]. Increasing the current density, thus the cathode field, enhances this positive charge generation. The same effect is observed for an AC stress. The magnitude of the shift for a given Q_{inj} depends on the charge injected per pulse (dose). The shorter the pulse, the sooner this shift appears (fig.2b). For a positive DC stress, V_{FB} goes through a minimum and then increases, which shows that. eventually, negative charge trapping takes over. This turnaround occurs for a higher Q_{inj} than previously reported for MOS capacitors on lightly doped substrates with the same oxide thickness [4]. For a negative DC stress, our data do not show a clear turnaround. If there is one at all, it occurs above 1 Cb/cm^2. For an AC stress, the position of the minimum depends on the charge injected per pulse. At

high negative current densities the turnaround seems to disappear.



a) positive DC injection

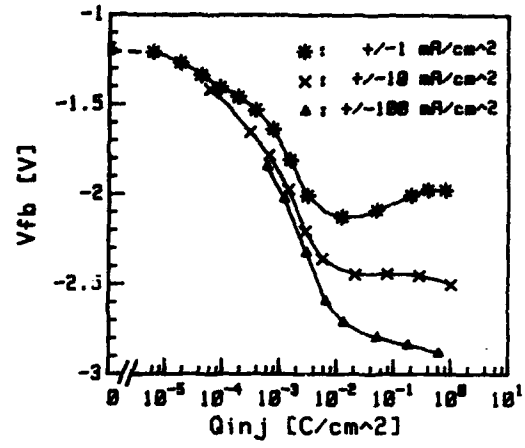


b) negative DC injection

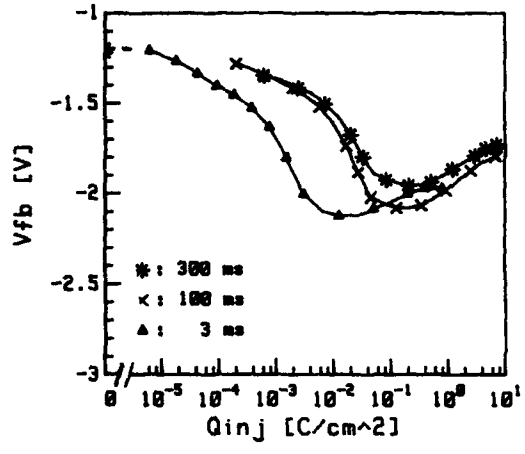
FIGURE 1
Flatband voltage as a function of injected charge for DC injection.

For comparing the results of C-V measurements with those obtained directly on EEPROM cells, it is more appropriate to vary the pulse parameters in such a way as to keep the dose constant. In fig.3, we separately varied the current density and duration of the pulses of each polarity. Positive charge generation is much more sensitive to the negative pulse, as could be expected from fig.1.

Measurements of V_{FB} under constant or pulsed voltage stress yield analogous conclusions.



a) constant pulse duration: 3ms

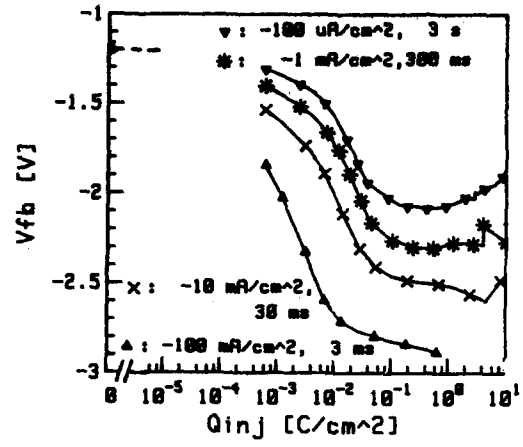


b) constant current density: +/-1mA/cm^2

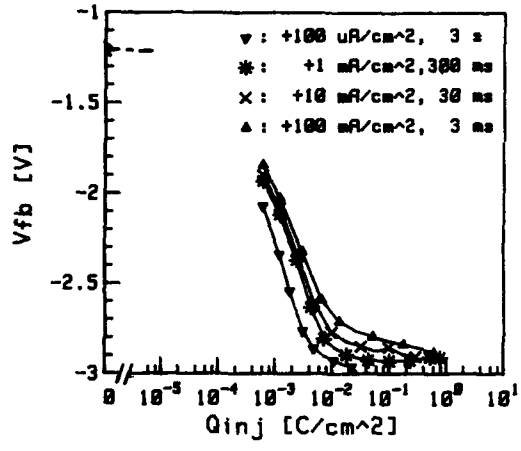
FIGURE 2
Flatband voltage as a function of injected charge for symmetrical AC injection.

The evolution of the memory window of EEPROM cells as a function of the number of programming cycles is shown in fig.4. The window initially o, 's (positive charge generation) and then closes (negative trapping). The former is most sensitive to the risetime, that is to the peak electric field [3], of the ERASE (negative) pulse. In contrast, negative charge trapping is only weakly affected by pulse risetimes.

In EEPROMs, the turnaround occurs between 10-100 cycles, which corresponds to about 10^{-4}-10^{-3} Cb/cm^2, since the charge injected per cycle is approximately 10^{-5} Cb/cm^2. In contrast, the lowest Q_{inj} for turnaround in our C-V measurements is



a) variable negative pulse

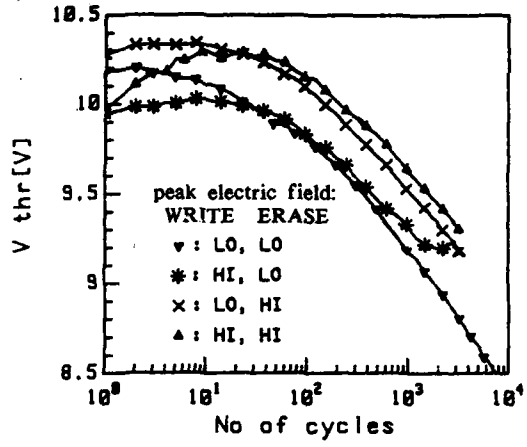


b) variable positive pulse

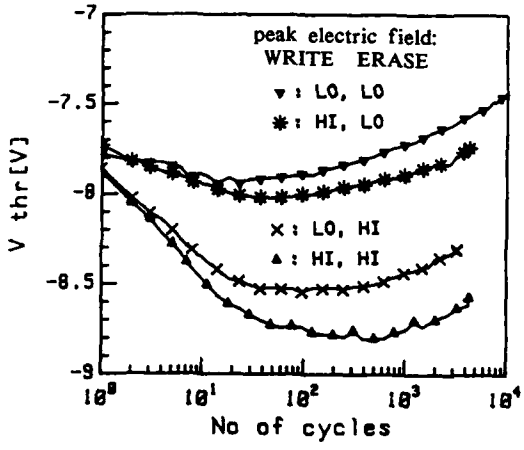
FIGURE 3
Flatband voltage as a function of injected charge for asymmetrical AC injection (constant charge per pulse = 300 uC/cm^2).

10^{-2} Cb/cm^2 (fig.2b). The discrepancy may be explained by the different doses and current levels used in the two measurements. In EEPROMs there are very likely high current spikes during WRITE and ERASE, as shown by calculations. Accurate values are difficult to estimate. Since the turnaround is shifted to smaller Q_{inj} for smaller doses, one can expect that for the conditions prevailing in EEPROMs, AC injection in MOS capacitors would give similar results. Instrumental limitations have prevented us from checking this conclusion.

Since window closing is more pronounced on the WRITTEN side (positive V_{th}), we conclude that negative charges must be closer to the substrate than to the floating gate.



a) WRITE threshold voltage



b) ERASE threshold voltage

FIGURE 4
EEPROM threshold voltages as a function of W/E cycling (W/E pulses: -22.5 V, 100 ms).

Similar results were obtained on devices fabricated with a different CMOS process and a 18nm thick injection oxide. In this case, the shift of V_{FB} was approximately twice as large. For EEPROM cells, the shift of V_{thE} was even more important.

4. CONCLUSIONS

Measurements on MOS capacitors and individual EEPROM cells show similar charge trapping behavior. Positive charge generation is most sensitive to negative bias on the gate. Négative charge trapping is not observed in our C-V measurements at high negative current density. Charge trapping under AC injection is not just a superposition of the contributions expected from DC injection measurements and reveals new features like dependence on the stress history (acceleration of positive charge trapping for smaller doses). Our results show that, with a careful choice of pulse parameters, one can optimize the wearout of EEPROM.

ACKNOWLEDGEMENTS

We wish to thank J.Solo de Zaldivar (Faselec) and H.Hofbauer (EM-Microelectronic) for providing samples.

REFERENCES

[1] P.Fazan. M.Dutoit, J.Manthey, M.Ilegems, J.M.Moret, ECS Fall Meeting, San Diego, Oct.19-24 (1986), Ext.Abstr.86-2, p.604.

[2] J.Solo de Zaldivar, NTG Fachberichte 77, 22 (1981).

[3] J.Manthey, M.Dutoit, M.Ilegems, ESSDERC'86, Cambridge, Sept.8-11 (1986)

[4] P.Fazan. M.Dutoit, C.Martin, M.Ilegems, Solid St.Electron., in print.

Degradation phenomena of tunnel oxide floating gate EEPROM devices.

J.S. Witters*, G. Groeseneken**, H.E. Maes.

IMEC v.z.w., Kapeldreef 75, 3030 Leuven, Belgium.

The degradation of tunnel oxide floating gate EEPROM devices was studied by using charge-pumping which allows direct characterization of the interface degradation on the transistor itself. It is found that positive charge is generated at the Si-SiO_2 interface, while negative charges are trapped at the injecting interface or in the bulk of the oxide. Combining these findings with a study of the influence of the charges present in the oxide and at the interfaces on injection and threshold voltage, it is possible to explain qualitatively all measured degradation characteristics.

1. Introduction.

High field stressing of thin oxides is known to cause charge trapping in the oxide and at the interfaces as well as generation of interface states [1,2]. The degradation is strongly dependent on the used stress conditions [3]. Although most of the experimental results have been obtained on capacitors, the same degradation behaviour is expected during high field stressing of MOS transistors. During stressing with alternating field polarities, as is the case for Electrically Erasable PROM (EEPROM) devices, the degradation mechanisms are not well understood until now [4].

In this work the degradation of tunnel oxide floating gate EEPROM devices is studied in detail by combining the charge pumping technique with threshold window degradation measurements. Contacted floating gate transistors are used to study the effect of Fowler-Nordheim injection on thin oxide gate dielectrics. Combining the measured results with a study of the influence of the charges present in the oxide and at the interfaces on injection current and threshold voltage, it is possible to explain all measured degradation characteristics.

2. Experiments and devices.

The studied EEPROM structures are n-channel floating gate devices in a 3μm p-well CMOS process with a thin oxide of 12 nm above the entire channel. In order to examine the effect of Fowler-Nordheim injection on thin oxide gate dielectrics, otherwise identical but contacted floating gate transistors were stressed with either positive or negative voltage at the floating gate while keeping drain, source and subtrate grounded. Subsequently they were characterized by means of the charge pumping technique.

The degradation of the EEPROM structures is studied by cycling the memory devices and monitoring the external threshold voltage (defined as the control gate voltage needed to allow a drain current of 1μA at a drain-source voltage of 2V). The fields applied during programming are of the order of 10 MV/cm. Erasing of the memory cell is achieved by Fowler-Nordheim injection of electrons towards the floating gate (uniformly over the whole channel area); writing can be accomplished either uniformly over the whole channel (by applying a negative voltage at the gate) or non-uniformly at the drain (by applying a positive voltage at the drain while keeping substrate and control gate grounded). Charge pumping was also performed on these memory devices . Indeed, a floating gate transistor can be considered as a MOS transistor with a stacked gate dielectric. The charge on the floating gate is equivalent to a uniformly distributed gate oxide charge. By applying the same formulas as used for conventional MOS transistors, we can calculate the density of interface states in floating gate devices with the same accuracy as for MOS transistors. The presence of non-uniformly distributed charges in the gate oxide can also be detected, as demonstrated previously [6].

3. Fowler-Nordheim injection in MOS transistors.

First thin oxide MOS transistors (contacted floating gate transistors) were examined. Fig. 1 shows the charge pumping measurements after different times of uniform

* Research assistant I.W.O.N.L. ; ** Senior Research Assistant N.F.W.O.

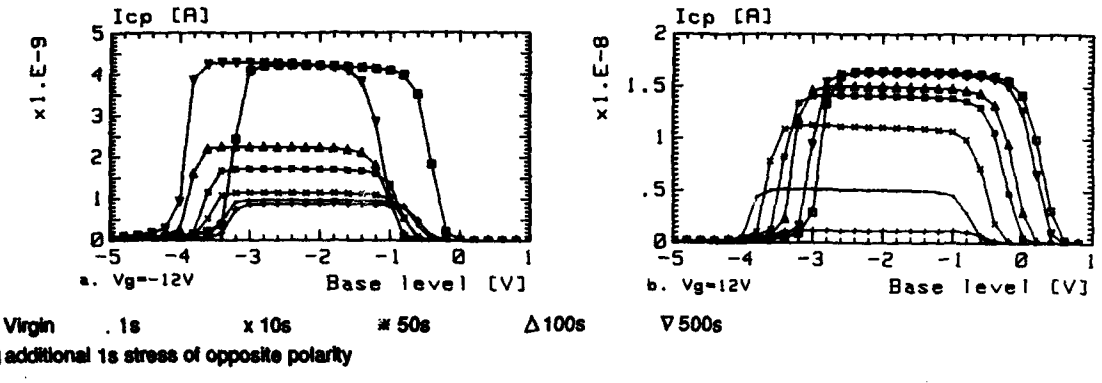


Figure 1. Charge pumping measurements after different stress times of uniform Fowler-Nordheim injection.

Fowler-Nordheim injection over the entire channel. The increase in charge pumping current indicates an increase of the interface state density. The shift of the edges of the curves implies the presence of charges in the gate oxide or at the interfaces. Fig. 1 also shows the results of charge pumping measurements on the same structures after an additional 1 second stress of the opposite polarity.

We can conclude from fig. 1a that Fowler-Nordheim injection of electrons from the gate (Vg<0) generates interface states and a positive charge in the bulk of the oxide or at the Si-SiO_2 interface (as detected by charge pumping). The positive charge can easily be removed by a short injection from the other electrode. It was also noticed that the positive charge disappears rather quickly under zero field conditions. This could mean that the charges are trapped very close to the interface. Longtime injection of electrons towards the gate (Vg>0) generates interface states. In this case however, the net trapped charge, as seen on fig. 1b is at first positive but then becomes negative with increasing stress time, consistent with [7]. A short pulse of the opposite sign, after the negative charge build-up, has little influence. A comparison of the effectiveness of interface state generation under both injection conditions is difficult since the experiments are done at equal stress voltages and thus not necessarily at equal stress currents.

4. Charges trapped in floating gate devices.

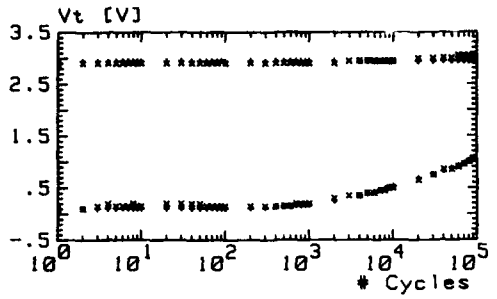
The external threshold voltage of a floating gate transistor is determined by the amount of charge stored on the floating gate and by the intrinsic threshold voltage of the transistor, defined as the external threshold voltage with no charge on the floating gate. The degradation of the threshold window is therefor caused by the combined effect of the changes of both.

Charges that are trapped in the oxide underneath the floating gate are changing the intrinsic threshold voltage. This charge also causes a variation of the injection field at the Si-SiO_2 interface, which influences the Fowler-Nordheim tunneling injection from the substrate (if the charges are located not to close to the interface [8]).

E.g. a negative charge located in the thin oxide will reduce the Fowler-Nordheim current towards the floating gate during erasing. Therefore, the amount of electrons stored on the floating gate after the programming operation, will diminish causing a decrease of the external threshold voltage. But the negative trapped charge will cause the intrinsic threshold voltage to increase. It can be shown that the influence of both changes on the external threshold voltage match exactly. We can conclude that uniformly distributed trapped charges (positive or negative) in the thin oxide have no effect at all on the external threshold voltage after an erase operation of a memory cell.

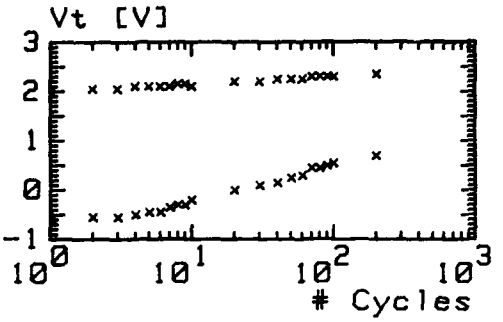
If writing is accomplished uniformly over the entire channel area, a negative trapped charge will reduce the tunnel current. So, after the write operation, less electrons will have been transported from the floating gate towards the substrate, thus leaving the floating gate at a more negative potential than in the case of no trapped charge. At the same time the trapped negative charge will increase the intrinsic threshold voltage. In this case, both effects will result in a more positive threshold voltage. A simple calculation shows that the combined effect is dependent only upon the magnitude of the trapped charges and is independent of the centroid in the thin oxide.

If the writing is done non-uniformly, the trapped charges will be distributed non-uniformly as well. In this case, one has to consider seperatly the injection region, which determines the



x measurement + simulation
Erase : Vg=15V, 10ms. Write : Vg=-16V, 10ms.

Figure 2. Degradation characteristic for uniform writing.



Erase :Vg=13V, 0.9s. Write : Vd=10V, 0.8s.

Figure 3. Degradation characteristic for non-uniform writing at low drain voltage.

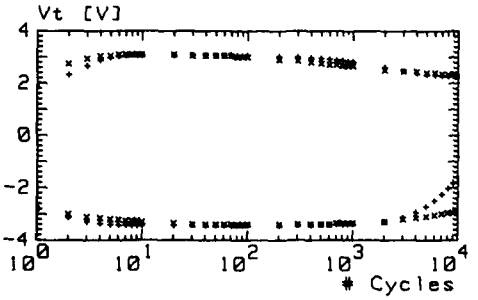
the amount of charge stored on the floating gate, and the channel region, which determines the intrinsic threshold voltage. As demonstrated further on, a non-uniform write operation can introduce a large amount of positive charge at the drain side. The largest part of the channel will not be affected by this charge. So, the tunnel injection current will increase due to the positive charge at the drain, but the intrinsic threshold voltage will not be altered, leading to a net decrease in the external threshold voltage after writing. If the write operation is now followed by an erase operation, the measured threshold voltage after erasing will be increased due to the same enhanced injection current effects at the drain.

5 . Degradation characteristics.

As indicated above, two different programming modes were used. For a uniform writing operation, the degradation proceeds straightforwardly : all trapped charges are distributed uniformly over the channel. The degradation characteristics shown in fig. 2 can easily be explained by considering the filling of electron traps following a simple first order kinetic model and by assuming the generation of electron traps during Fowler-Nordheim injection, as reported before [10]. Simulation results are also shown in fig. 2. Remarkable is that none of the measurements shows the presence of a positive charge during the first cycles : the alternating high fields do not seem to allow the permanent trapping of a positive charge, in contrast with the dc stress results of fig. 1.

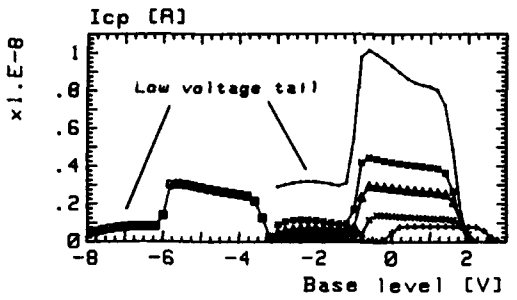
Degradation under non-uniform writing conditions is a more complicated process. If the applied drain voltage is fairly low, the degradation is governed by electron trapping only. The observed behaviour is roughly the same as for uniform degradation (fig. 3). The process is faster due to the larger fluence (electrons/unit area) through the thin oxide at the drain side. The observed degradation is mainly caused by a decreased injection from the floating gate. The positive threshold voltage increases slowly, possibly due to the effect of the negative charge at the drain side and the interface states on the transistor characteristics combined with some negative charges trapped very close to the Si-SiO2 interface in the channel area, having a dominant effect on the intrinsic threshold voltage and a rather weak influence on the current injection [8].

If the drain voltage is higher such that a hole current into the substrate is generated due to the deep depletion region in the drain [9], the degradation behaviour is totally different. We observe trapping of positive charge in the thin oxide at the drain during the write operation, causing the opening of the threshold window (fig. 4). The charge is clearly identified by the low voltage tail in the charge pumping characteristic (fig. 5). With increasing number of cycles the window is closing again, but as the two threshold voltages do not start to degrade simultaneously, more than one mechanism must determine the window degradation. Trapping at the drain of negative charge can explain the closure of the threshold window but for the



x measurement + simulation
Erase : Vg=15V, 10ms. Write : Vd=15V, 8ms.

Figure 4. Degradation characteristic for non-uniform writing at high drain voltage.



+ 10 cycles x 2E3 cycles Δ 1E4 cycles
* 2E4 cycles . 1E5 cycles
□ 1E4 cycles at low threshold voltage

Figure 5. Charge pumping measurements during non-uniform writing after different numbers of cycles.

high threshold voltage starting to degrade earlier, the centroid of the negative charges should be located close to the Si-SiO2 interface. This is however in contradiction with the charge pumping measurements, where a positive charge tail remains visible. A plausible explanation is that due to the enhanced injection at the drain, the channel area remains in the 'initial' regime (fig. 1b and [7]) for a large number of cycles, during wich Fowler-Nordheim tunneling gradually generates positive charges. This positive charge has a decreasing influence on both the high and the low threshold voltage. The combined effect of positive and negative charges at the drain and a positive charge in the channel region, can account for the whole degradation behaviour. Simulation results, based on these assumptions, are also depicted in fig. 5 ; they are showing the same qualitative behaviour as the measurements.

6. Conclusions.

An analysis of the degradation behaviour of floating gate devices must be based on a careful study of the influence of trapped charges in the floating gate transistor. When the write operation is performed non-uniformly, an important phenomenon is the substrate current due to the deep depletion in the drain [9]. If no substrate current is present during programming, no window opening has been observed.

REFERENCES

[1] C.Hu, Techn. Dig. of 1985 IEDM, p.368-371.
[2] L.DoTanh, P.Balk, Proc. of Insulating Film on Semiconductors, 1983, p16-19.
[3] M.Liang, S.Haddad, W.Cox, S.Cagnina, Techn Dig. of1986 IEDM, p.394-398.
[4] B.Euzent, N.Boruta, J.Lee, C.Jenq, IEEE Proc. IRPS 1981, p11-16.
[5] G.Groeseneken,H.Maes, N.Beltran, R.De Keersmaecker, IEEE Trans. on Electr. Dev., vol 31, 1984, p42-53.
[6] P.Heremans,G.Groeseneken,H.E.Maes,presented at the 1986 SISC, San Diego.
[7] Z.A.Weinberg,M.V.Fischetti,Y.Nissan-Cohen, J. of Appl. Phys., vol 59, p 824-832.
[8] P.Solomon, J. of Appl. Phys., vol 48, p3843-3849.
[9] A.Kolodny,S.T.K.Nieh,B.Eitan,J.Shappir, IEEE Trans. on Electr. Dev., vol 33, 1986, p835-844.
[10] M.Heyns, R.DeKeersmaecker, M.W.Hillen, Appl. Phys. Letters, vol. 44(2), 1984, p202.

Session B1.3

Monte Carlo Device Simulation II

Chairman: B. Riccò

Monday, September 14, 1987

A MONTE CARLO APPROACH TO THE STUDY OF THE DRIFT-DIFFUSION TRANSPORT MODEL

Cinzia MANTILLI, Franco VENTURI, Bruno RICCÒ, and Enrico SANGIORGI

Department of Electronics, University of Bologna
Viale Risorgimento 2, 40136 Bologna, Italy

In this work, the Monte-Carlo simulation is used as a tool to investigate the field dependence of electron mobility and velocity saturation in MOSFETs. A simple, analytical model, in good agreement with experimental results as well as simulations, is studied from straightforward physical considerations.

1. INTRODUCTION

As known, the Monte-Carlo (MC) approach to simulation of current transport in semiconductors is more general than that based on the drift-diffusion (DD) model. Therefore in the (important) subset of cases where both are equally applicable it can be used as a tool to study crucial DD parameters that can be independently derived from suitable average of individual carrier characteristcs. In this sense, MC simulation represents an ideal complement of the comparison with experiments as it allows to independently evaluate single effects that, in reality, occur simultaneously and cannot be simply separated out.

In the present paper, this philosophy has been applied to the study of the field dependence of the carrier mobility (μ) and velocity saturation (v_{SAT}) in MOSFETs that, as known, represents a crucial problem in DD numerical simulations.

For clarity we will separately consider the dependence on the longitudinal and trasversal components of the applied field (E_L and E_T respectively).

2. LONGITUDINAL FIELD DEPENDENCE

In order to understand the essential features of the microscopic physics responsible for the decrease in electron mobility, we have exploited the unique characteristics of the MC approach by looking at the effects of the various scattering mechanisms among which, under the operating condition of interest, emission of optical phonons plays a dominant role. Since, when dealing with average properties of the electron gas, only relatively small energies need to be considered (typically up to a few hundreds of mV), the semiconductor band can assumed to be parabolic and the electron scattering rate with optical phonons (f_{EOP}) is then simply given by [1]

$$f_{EOP}(E) \propto \sqrt{E - E_{OPT}} \qquad (1)$$

(where E and E_{OPT} denote the electron energy and the threshold for the emission of optical phonons respectively).

An interesting point emerges considering the macroscopic average P_{EOP} of f_{EOP} and plotting $1/P_{EOP}$ and μ calculated in the same points as function of E_L.

Fig. 1 represents a typical set of MC results obtained for a 1D $n^+ - n - n^+$ structure with the uniform dopings (10^{20} and 10^{14} cm^{-3} for the n^+ and n regions respectively) with the simulator MOS^2 [2]. Several points along the lightly doped region have been studied to make sure that the results are independent of (specific) geometrical and technological characteristics. In these calculations μ has been obtained taking the ratio between the averaged carrier velocity and E_L, P_{EOP}, instead, simply counting the number of optical phonons emitted per unit time and unit volume.

The evident parallelism of the two curves namely

$$\mu \propto \frac{1}{P_{EOP}}, \qquad (2)$$

clearly indicates the mobility degradation to be essentially due to the larger emission of optical phonons resulting from the field induced increase of electron energy.

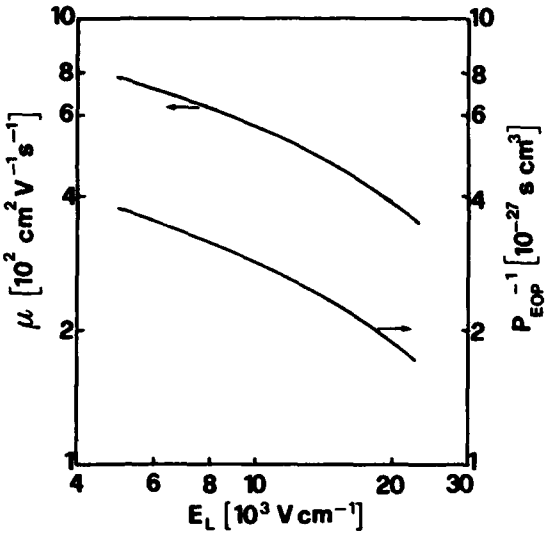


Fig.1 : μ and $1/P_{EOP}$ calculated with the MC simulator as a function of E_L in a central point of a $n^+ - n - n^+$ structure (well outside the space charge regions of the $n^+ - n$ junctions). The doping of the central portion of the simulated device is 10^{14} cm^{-3}.

Eq. 2 is a clear manifestation of the validity of the relaxion-time approximation of the transport equations (at least under the conditions of interest for real devices) from which the mobility dependence on the driving field can be worked out.

The well known model for the mobility dependence on the longitudinal field (E_L) in MOSFETs

$$\mu = \frac{\mu_{P0}}{\sqrt{(1 + (\mu_{P0} E_L / v_{SAT})^2}}, \qquad (3)$$

often considered to be only empirical in the device simulation literature, can then be analytically derived [3].

3. NORMAL FIELD DEPENDENCE

Although under the conditions of interest, optical phonons provide the dominant scattering mechanism, that with the $Si - SiO_2$ interface must also be considered to account for the mobility dependence on the transversal component of the applied field.

If P_I denotes the probability of such a phenomenon, extending the approach of the previous section we have

$$\mu_{P0} \propto (P_0 + P_I)^{-1} = P_0^{-1}(1 + P_I/P_0)^{-1} \qquad (4)$$

where P_0 represents the sum of the probabilities of other scattering mechanisms that varying only with the carrier energy are essentially independent on E_T . (The consideration of these other effects is here required if μ_{P0} has to mantain finite values for vanishingly small transversal fields).

The mobility dependence on E_T is then simply described by the factor $(1 + P_I/P_0)^{-1}$ whose accuracy could, in principle, be easily studied by means of MC simulations of devices biased with very small E_L so as to deal only with the degradation due to E_T. Unfortunately, small driving fields make the calculations extremely long, hence prohibitively expensive.

A much more practicable procedure consists of directly evaluating the analytical expression resulting from the assumed approach. The model for the (elastic) scattering with the interface included in MOS^2 is that described in Ref. [4,5] that has been favourably compared with others [6,7] looking at their effects on the electron distribution within the MOSFET channel and assuming the result of the DD simulation as a reference.

According with the chosen model, $P_I \propto E_{TS}^2$ (E_{TS} denoting the transversal field at the $Si - SiO_2$ interface) and eq. 4 finally gives

$$\mu_{P0} = \frac{\mu_0}{1 + (E_{TS}/E_C)^2} \qquad (5)$$

where μ_0 represents the low field carrier mobility and E_C denotes a (constant) critical field to be essentially determined by fitting experimental data.

With regard to eq.5 it is worth stressing that it features only the value of E_T at the $Si - SiO_2$ interface, thus describing a sort of global, rather then purely local effect. In practice this turns out to be particularly suitable for the simulation of MOSFETs where the overwhelming part of the carriers is close enough to the interface and E_{TS}, depending on the global charge within any given transversal portion of the device, is

much more accurately defined than the local E_T, often strongly dependent on the details of the mesh used to descretize the simulation domain.

Fig. 2 shows a comparison between the model of eqs. 3 and 5 and a well known set of experimental results taken from the literature [8]. As can be seen a good agreement is obtained with reasonable values of the adjustable parameters (in particular $\mu_0 = 950\ cm^2V^{-1}s^{-1}$, $v_{SAT} = 9.510^6\ cms^{-1}$, $E_C = 10^5\ Vcm^{-1}$).

The figure also contains a few selected results of complete MC simulations of MOSFETs to show the accuracy of the calculations used as a basis for the present work.

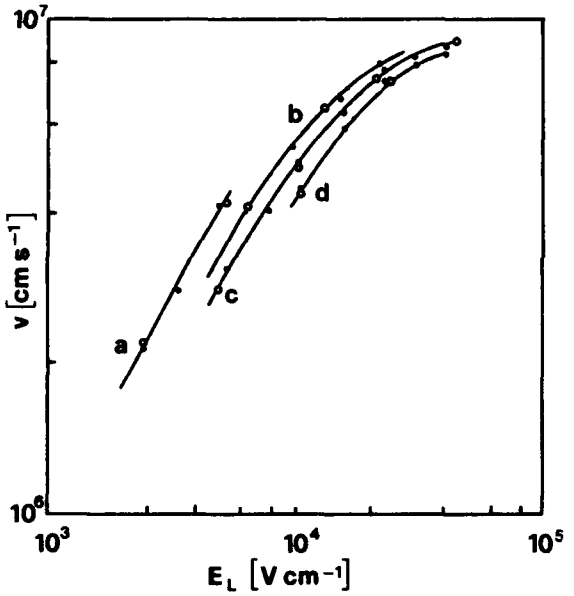


Fig.2 : Electron velocity vs. longitudinal field for the following values of transversal field: a)$E_T = 3 \cdot 10^4$ V/cm, b)$E_T = 6 \cdot 10^4$ V/cm, c)$E_T = 8 \cdot 10^4$ V/cm, d)$E_T = 10^5$ V/cm. (•) represent experimental data taken from the literature [8], (○) indicate the results of selected MC simulations and the solid lines are obtained from the analitical model of eq.3.

A more complete set of MC simulations compared with experimental data [9] is shown in fig.3.

Fig.4, instead, shows the dependence of the electron saturation velocity on E_T that is found to be substantial in the range of electric field of interest for submicron-size devices. From a physical point of view,

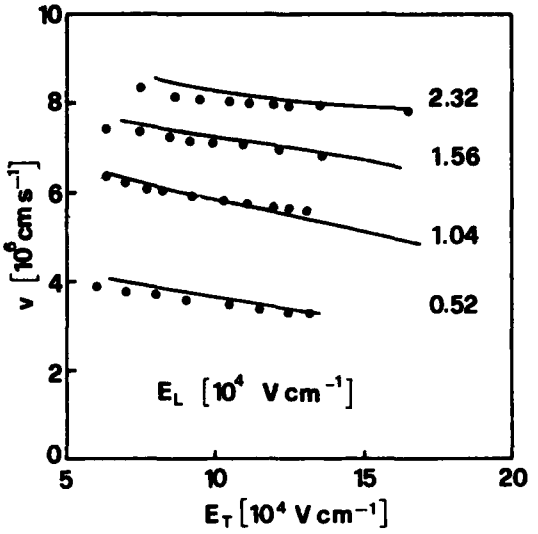


Fig.3 : Velocity vs. transversal field (E_T) for different longitudinal field values: the dots are experimental points [9] and the lines are obtained from MC simulations.

the effect of E_T on v_{SAT} is ultimatly due to the variation of the carrier interaction among themselves and with the $Si-SiO_2$ interface that depends on the actual current path and charge distribution within the device, thus exhibiting an intrinsically global two dimensional dependence on the applied voltage.

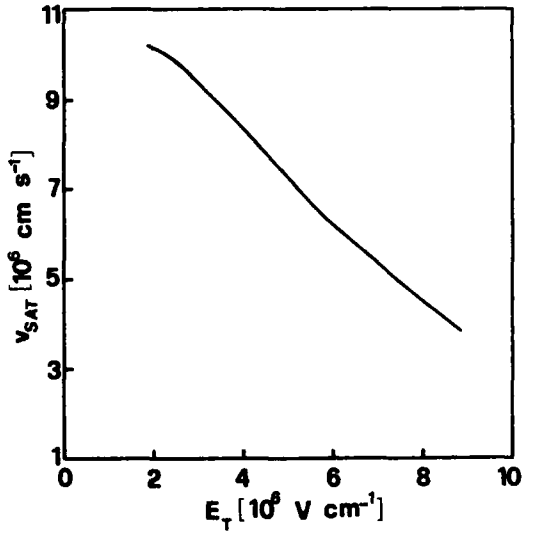


Fig.4 : Velocity saturation vs. transversal field E_T.

4. CONCLUSIONS

We have used Monte Carlo simulation of electron transport in MOSFETs to investigate the dependence of the carrier mobility on the applied field and the effect on the electron saturation velocity.

In this way the physical mechanisms responsible for mobility degradation and field dependence of v_{SAT} can be studied down at the microscopic level.

REFERENCES

[1] C. Jacoboni, L. Reggiani, " The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials ", Rev. of Mod. Phys., vol. 55, pp. 645-675, 1983.

[2] B. Riccò, E. Sangiorgi, F. Venturi, P. Lugli, " Monte Carlo modeling for hot electron gate current in MOSFETs ", IEDM Tech. Dig., pp. 559-562, 1986.

[3] E.M. Conwell, " High field transport in semiconductors", Solid State Physics suppl.9, pp. 187, 1967.

[4] D.K. Ferry, " The transport of electrons in quantized inversion and accumulation layers in III-V compounds ", Thin Solid Films 56, pp. 243-252, 1979.

[5] Chu-Hao, J. Zimmermann, M. Charef, R. Fauquembergue, E. Constant, " Monte Carlo study of two dimensional electron gas transport in Si-MOS devices ", Sol. Sta. Elec. vol. 28, pp.733-740, 1985.

[6] Y.park, T. Tang, D.H. Navon, "Monte Carlo surface scattering simulation in MOSFET structures", IEEE TED vol ED-30, pp.1110-1116, 1983.

[7] P.J. Price, "Monte Carlo calculation of electron transports in solids", Semiconductors and Semimetals, vol.14 pp.249-308, 1979

[8] J.A. Cooper, Jr. and D.F. Nelson, " Measurement of the high field drift velocity of electrons in inversion layers on silicon ", IEEE EDL, vol.EDL-2 171, July 1981.

[9] J.A. Cooper, Jr. and D.F. Nelson, " High-field drift velocity of electrons at the as determined by a time-of-flight technique", J.Appl.Phys. 54(3), March 1983.

HOT ELECTRON DYNAMICS MONTE CARLO SIMULATION IN HETEROSTRUCTURE SEMICONDUCTOR DEVICES

Francesco ANTONELLI *

IBM European Center For Scientific and Computing
Via del Giorgione 159, 00147 Roma, Italia

Paolo LUGLI

Dipartimento di Fisica, Università di Modena
Via Campi 213/A, 41100 Modena, Italia

This paper reports on a Monte Carlo simulation of the hot electron transport phenomenon in an heterostructure semiconductor device. Two different electron populations have been simulated : the hot electrons injected via a tunneling mechanism into the base, and the thermal electrons arising from the high doping density of the ballistic device. Electron-electron scattering and plasmon- electron scattering have been introduced into the physical model which includes also electron degeneration and quantum reflections at the collector barrier. The simulation has been compared with the experimental results obtained from the THETA device at $4.2K$.

1. INTRODUCTION

One of the most outstanding trends in semiconductor devices technology consists in the reduction of the switching time. As a result of this operation it is possible to obtain high computing velocity, improving computer performances independently from its architecture. The most common technology used in the fastest switching logic circuits is the MESFET technology (Metal Semiconductor Field Effect Transistor) using, as semiconductor material, Gallium Arsenide (GaAs). Within this class of devices it is possible to achieve switching times of the order of $50ps$. Further improvements have been obtained with the HEMT (High Electron Mobility Transistor) technology which is capable of switching velocity of the order of $10ps$. Very high speeds are obtained by minimizing the energy and momentum losses that electrons undergo as their cross the active area of the device. In this direction, a very promising possibility resides in the concept of Hot Ballistic Electrons; in a standard device, as MESFET, and also in HEMT, electrons are injected into the channel with a thermal energy distribution and a small initial velocity. In order to reduce the transit time through the channel or the base, it is very important to increase their initial velocity; this is accomplished in a class of devices that we will call HEIBD [1] (Hot Electron Injected Ballistic Device). In HEIBD, electrons are injected into the base with energies of some hundreds me$V$ greater than the thermal energy; this technique promises switching time of the order of $1ps$. Looking at these estimates,it is to recognize the importance of a systematic study of the properties of ballistic semiconductor device, and in particular of HEIBDs. We present here a Monte Carlo simulation of a device called THETA (Tunneling Hot Electron Transfer Amplifier) [2], but the model applies also to other classes of ballistic devices. The device (see Fig.1a), is a heterostructure consisting of alternate layers of GaAs and AlGaAs; electrons are injected into the base via tunnel effect through the potential barrier between emitter and base. Because of the homogeneity of the electric field in the normal direction to the plane of Fig. 1a, it is possible to restrict ourselves to a two dimensional real space simulation.

The main aspect of this work is that, in addition to the electron degeneration effects, the quantum structure of the potential barrier at the base collector interface has been taken into account; this is an important point in order to compare the model predictions with experimental data. All the computations have been performed on an IBM 3090-200/VF, which is particularly suitable becuse it is a vector-parallel system, and

because it has a large central storage capability.

In section 2 we will describe the physical model that has been used while section 3 and 4 will be devoted to the discussion of results and conclusion.

2. THE PHYSICAL MODEL

A two valley GaAs model is used for the simulation. The scattering mechanisms considered are with polar optical phonons, non-equivalent (Γ-L) and equivalent (L-L) intervalley phonons, acoustic phonons, and ionised impurity. The Monte Carlo algorithm is coupled to a Poisson solver in the way described in Refs. 3 and 4. In HEIBD the quantum transport properties are strongly affected by the shape of the conduction band edge which is shown, for the particular device that has been studied, in Fig. 1b. This structure and the high doping density, which in the THETA device is of the order of $10^{24}/m^3$, introduced in order to obtain reasonable emitter currents with a bias smaller than the separation between Γ and L valleys, impose several quantum corrections necessary for an accurate description of HEIBD dynamics. The quantum effects

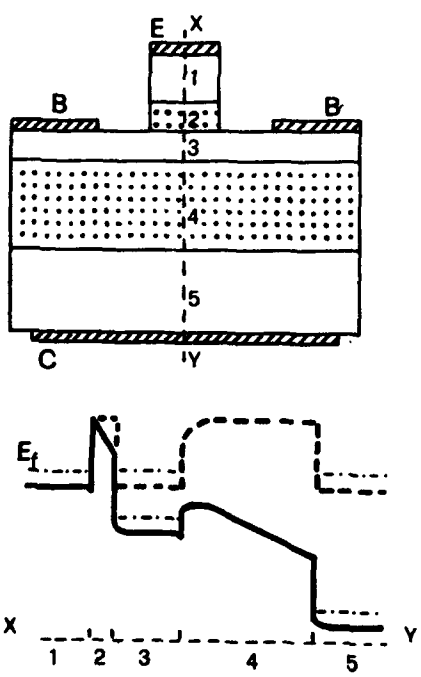


Figure 1: (a) schematic representation of the simulated device. The dotted regions (2) and (4) are made of AlGaAs, the others of GaAs; (b) conduction band edge without (dashed curve) and with bias (solid curve).

are included in scattering rates (through Pauli exclusion principle), in distribution functions and in transmission and reflection coefficients.

Some other important feature induced by the peculiar structure of HEIBD are listed below:

a) The doping in the base region produces a "sea" of cold degenerate electrons (in the simulated device about 10^5 electrons.) These electrons are confined in the base by the collector potential barrier; the presence of such electron "plasma" is an important energy or longitudinal momentum loss mechanism both on short and long range. Electron-electron and electron-plasmon scattering are considered as in Ref. 5. These two scatterings are implemented splitting the Coulomb Potential in a long range part, responsable for the electron-plasmon scattering, and a short range part ,responsible for the electron-electron scattering. The screening length, discriminating between the two different regimes, is the Fermi-Thomas length. To avoid unphysical long range correlations in the electron-electron scattering the partners are chosen, if there are any, in the range of two Fermi-Thomas lengths.

To account for the Pauli exclusion principle which affects remarkably the dynamics of the highly degenerate particles considered, we used the technique which the k-space is divided in elementary cells of arbitrary volume, the occupation number (i.e. the maximum number of particles contained in agreement with Pauli exclusion principle) is obtained by counting the number of states contained in each cell (see Refs. 3 and 6 for details). In order to consider particle inhomogeneities in the base, four different occupation numbers are computed in terms of the dimension of the region and assigned to a particular k-space matrix. Figure 2 shows the particle distribution in the base GaAs layer after 100 time step and corresponds to region 3 of Figure 1a; the depletion near the emitter *AlGaAs* interface



Figure 2: Distribution of particles in the base. The four regions in the simulation are divided by the dashed lines.

is visible even if it is partially filled by some injected electrons. Two depletions are recognisable also near the base electrodes.

b) The second relevant point is the presence of the emitter and collector barriers. The spectrum of injected electrons is obtained integrating numerically the Schroedinger equation for incident plane wave functions. In this way one obtains the transmission coefficient as a function of total electron energy $T_e(E)$ and normal electron energy $T_e(E_n)$, where E_n is the energy associated with momentum component k_n normal to the collector barrier. Two approximations have been used: no alloy scatterings have been considered, such that the tunneling is perfectly elastic, the influence of the collector barrier has not been taken into account (resonant tunneling). Figure 3a shows the distribution of injected particles $N(E_n)$ at 4.2 K which, in terms of current density distribution, is given by:

$$N(E_n) = \int_0^T \int_S J(E_n) ds dt \qquad (1)$$

and $J(E_n)$ is current distribution per unit surface:

$$J(E_n) = \frac{m_e}{2\pi\hbar^3} T(E_n) \int_0^\infty f_E(E)[1 - f_B(E)] dE_t \quad (2)$$

where f_E and f_B are respectively the density of states in the emitter region and in the base region, and E_t is the energy associated with the momentum components parallel to the collector barrier.

c) The last to analyse is the effect of a quantum collector barrier. As already mentioned, the resonant tunneling has not been included in the model, thus it was possible to consider separately the collector barrier and compute numerically the transmission and reflection coefficients $T_c(E_n)$ and $R_c(E_n)$. It is very likely that resonant tunneling has a considerable influence on HEIBD dynamics, but in any case the presence, in connection with the other effects previously described, of the quantum collector barrier, not considered in other simulation, permits a clearer interpretation of experimental results.

3. RESULTS

We present here some of the results of the simulation, compared with experimental data. A more complete discussion can be found in Ref. 7. Figure 3b represent the collected electron spectrum for an emitter current of $100\mu A$ and a collector barrier eight of 170 meV. A comparison with the injected particle spectrum reported in Figure 3a shows that the width of the distribution at half maximum, of about 65 meV, is of the same order of the injected electron distribution, and this is a clear signal of the ballistic behaviour of the injected electrons.

For the analysis of the characteristics curve few preliminary comments are needed. For a graded collector barrier as in the THETA device the dependence between collector bias ΔV_c and barrier height ΔE_c is linear only for negative bias values and, in this case, is given by:

$$\Delta E_c = \Phi_c - \Delta V_c \qquad (3)$$

where Φ_c is the collector barrier height for zero bias. For $\Delta V_c > 0$ this relation is no longer valid and is strictly connected with the grading profile which is, at present, not very well know. For Φ_c we used the value commonly quoted in the literature of $215^{+35}_{-15} meV$, where the errors are related to the uncertainty on the Al mole fraction in AlGaAs. The insert in Fig. 4 shows the theoretical energy distribution obtained using the linear approximation given by Eq. (3) (circles), compared with experimental result (solid line). In this case one has to consider that the experimental curve is obtained indirectly, deriving the (I,V) diagram with respect to the collector bias. The two points in the interval (-1.0,0.0) V are due to an enhancement originated by the

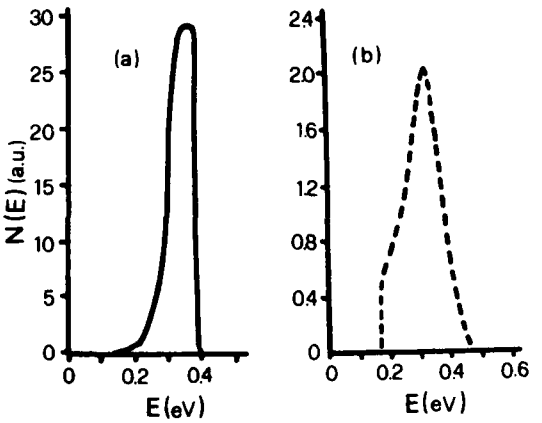


Figure 3: (a) Energy spectrum of the injected particles for $I_e = 100\mu A$; (b) energy spectrum of the collected particles, with a collector barrier height of 170 meV.

phonon interaction and are not present in the experimental curve, probably because the real injected particle distribution is affected by interbarrier diffusion, inelastic tunneling should then be considered. Analysing in more detail the (I,V) curve one can see the small tail in the interval (-0.3,-0.2) V due to the corresponding tail in the electron distribution. The degeneracy of the electrons is now fundamental preventing about 85As previously mentioned the comparison of the data is not straightforward for the positive collector bias because of the presence of the grading at the collector barrier. In this case it is not possible to use the linear relation (6); in fact the uncovered region between 0.0 and 1 V is actually still under investigation, in particular for what concerns the grading profile. Using, at 1 V bias, the value for the collector barrier of 165 me$V$ reported in the literature, we obtain values for the transfer ratio in good agreement with experimental results. In the same hypotheses Figure 4, where we reported all Monte Carlo results for $I_e = 100\mu A$ (circles) and $I_e = 60\mu A$ (crosses), shows that the point obtained for 1.0V base-collector bias is in good agreement with experiment (solid line). Then, assuming that the value of 165 me$V$ is correct, this point is a good check of the model because in this case the effect of the resonant tunneling is quite irrelevant.

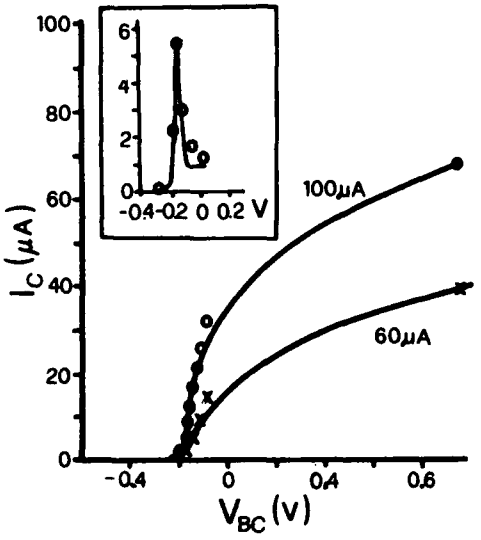


Figure 4: Collector current for $I_e = 100\mu A$ and $I_e = 60\mu A$. The solid line shows the experimental curves, circles and crosses the Monte Carlo results. The insert displays the comparison between the theoretical and experimental energy distributions.

4. CONCLUSIONS

In conclusion, we have shown that it is possible to study ballistic transport using a semiclassical approximation including very careful and sophisticate treatment of quantum effects; the data reported, deprived from arbitrary derivation from the energy distribution show that the state of art is skillful enough to obtain a good description of HEIBD dynamics. The model used is capable also of successfully describe other ballistic devices. Furthermore,an analysis of the computing time, for the code that we used in the present work, suggests that the introduction of the next level of complexity (including a more complete evaluation of the band structure and of the equation of motion) is possible with the present generation of computers. The problem of resonant tunneling , which seems to play an important role in the dynamics of HEIBD, need additional investigations.

ACKNOWLEDGEMENTS

The authors want to thank IBM for the computational facilities at Rome European Center for Engineering and Scientific Computing (ECSEC). Drs. P. Sguazzero, R.W. Hockney and C. Jacoboni are also acknowledged for suggestions and very useful criticisms.

REFERENCES

1. J.R. Hayes, A.F.J. Levi and W. Wiegmann, *El.Lett.* **20** (1984) p.851.
2. M. Heilblum, *Sol.Stat.El.* **24** (1981) p.343.
3. P. Lugli and C.Jacoboni, these proceedings.
4. P.Lugli and D.K. Ferry, *Physica* **134B** (1985) p.364.
5. S. Bosi and C. Jacoboni, *J. Phys.* **C9** (1976) p.315.
6. R.W. Hockney and J.W. Eastwood, *Computer Simulation Using Particles* Mc Graw Hill (1981).
7. F. Antonelli and P.Lugli, submitted for pubblication to IEEE Trans.on Electron Dev.

*Present address: Honeywell Bull Italia, Via A. Danoli 2, Firenze, Italia

A MONTE CARLO ANALYSIS OF DIFFUSION-NOISE PROPERTIES IN GAAS-ALGAAS QUANTUM WELLS

Rossella BRUNETTI

Dipartimento di Fisica, Università di Modena
Via Campi 213/A, 41100 Modena, Italia

Steven M. GOODNICK

Department of Electrical and Computer Engineering
Oregon State University, Corvallis, Oregon, U.S.A.

We present a general analysis of steady-state velocity fluctuations, diffusion and noise for electrons in GaAs-AlGaAs quantum wells under high-field conditions using an Ensemble Monte Carlo simulation. Here we analyze for the first time diffusivity and noise problems in a two-dimensional structure by means of the velocity autocorrelation function. From the Monte Carlo simulation we obtain results for the diffusion coefficient at various field strengths and for the autocorrelation function of velocity fluctuations as a function of time for different physical conditions. We also show the power spectral density of velocity fluctuations as a function of frequency for such systems, and a comparison is made with the available data in the literature. The role of interband scattering, a new noise source not present in bulk structures, is discussed, together with the comparison between bulk and 2D results.

1. INTRODUCTION

A theoretical analysis of fluctuations and noise in semiconductors can yield relevant information on the microscopic interpretation of transport coefficients as well as on the detailed features of various scattering mechanisms. In submicron devices, where high fields are present and ultrafast transport processes are usually involved noise and diffusion may play an important role in the design and characterization of the device itself. At present very little experimental data exists on the noise properties of quantum wells and few theoretical investigations have been made on fluctuation phenomena in 2D systems [1, 2].

We treat the theoretical problem in terms of the autocorrelation function, a quantity directly related to diffusivity and noise, and we obtain results for realistic structures using an Ensemble Monte Carlo simulation of electrons confined to a quantum well.

2. THEORY AND THE NUMERICAL PROCEDURE

Let us consider an ensemble of non-interacting electrons in a semiconductor subject to an external electric field **E** and to the action of scattering agents (phonons, impurities,-etc). If the electrons are nonuniformly distributed in space the phenomenon of diffusion occurs, tending to uniformly spread the concentration of carriers in space.

This behaviour is described, at a phenomenological level, by the diffusion equation [3]:

$$J_i = e\{n(\mathbf{r})v_{di} - D_{ij}[\frac{\partial n(\mathbf{r})}{\partial x_j}]\} \qquad (1)$$

In the above equation e is the electron charge, r is the spacial position with components x_i, $n(\mathbf{r})$ and J the particle density and current density, respectively, D_{ij} is the diffusion coefficient tensor (i,j=1,2,3, the sum over repeated indices is implied), $\mathbf{v_d}$ is the drift velocity of the carriers in absence of diffusion.

If **E** is applied along a high symmetry direction of a cubic crystal, then D_{ij} reduces to a diagonal form, with a longitudinal component D_l and two transverse components D_t. In the following we will analise diffusivity and noise properties along the field direction; consequently we will use a simplified scalar notation where all the vector quantities are substituted with their longitudinal components.

For vanishingly small electric-field strengths, diffusivity D and mobility μ are field independent and satisfy the Einstein relation [3]

At high fields the Einstein relation fails and the study of diffusion is generally performed through the introduction

of a field-dependent D [4].

If the concentration gradients are small and in absence of carrier-carrier interaction $D(\mathbf{E})$ can be obtained from the following equation [3]:

$$\frac{d\langle(x-\langle x\rangle)^2\rangle}{dt} = 2D \qquad (2)$$

where x is the displacement along the field direction and brackets represent an ensemble average. The quantity on the right-hand side is the second central moment (SCM) of the distribution $n(x)$; Eq.(2) is valid at times longer than both the transient transport time and the correlation time, i.e. the time necessary for setting up the correct space-velocity correlations which are at the basis of diffusivity effects [5].

The phenomena of diffusion is stricly related to noise. The mathematical quantity that describes the common origin of diffusion and noise is the autocorrelation function (ACF) of velocity fluctuations, which contains information on the magnitude of the fluctuations and how they decay in time:

$$C(t) = \langle \delta v(r)\delta v(r+t)\rangle \qquad (3)$$

(the mean value in steady state conditions is independent of r). $C(t)$ is related to the diffusion D through:

$$D = \int_0^\infty dt C(t) \qquad (4)$$

Thus D can be evaluated through the knowledge of C(t), which is of interest in itself since it gives important physical information on the time evolution of the carrier dynamics .

Finally, we introduce the noise spectrum $S_v(\omega)$:

$$S_v(\omega) = lim_{T\to\infty}\langle\left|\int_0^T \delta v(t)e^{i\omega t}dt\right|^2\rangle \qquad (5)$$

The noise spectrum is then related to the ACF C(t) by the Wiener-Kintchine theorem [6].

The 2D physical system under investigation is a square well representing the effective 1D potential arising from the band offset between GaAs and AlGaAs (mole fraction of Al in the alloy =.23). The solutions of the wave equation for this potential give rise to a series of 2D dimensional subbands which we use in calculating the scattering rates for electronic motion parallel to the well, modeled using an Ensemble Monte Carlo (EMC) simulation. For a well width of 100Å two subbands are allowed in the central valley. We treat both intra- and intersubband scattering of the 2D electrons by bulk longitudinal optical phonons [7]; intervalley scattering to the satellite L-valleys is also included. Details of the physical model and the numerical procedure can be found in Refs. [8] and [9]. We do not consider the effect of intervalley transfer of carriers out of the well in this work.

During the numerical simulation, we follow the time evolution of an ensemble of 4000 electrons whose free flights are generated stochastically according to the calculated scattering probabilities. From this ensemble, we calculate the SCM of the carrier displacement and the ACF of velocity fluctuations about their mean value.

D is determined both from the SCM following eq.(2) and from the ACF using eq.(4). $S_v(\omega)$ is calculated as the Fourier transform of the ACF from the Wiener Kintchine theorem. For the details of the numerical evaluation of the SCM and the ACF see Ref. [4].

In order to compare 2D results with 3D results an EMC program for bulk GaAs has also been used [4]. The physical model for GaAs includes the same intravalley and intervalley scattering sources as the 2D program, and the input parameters of the material have been consistently chosen.

3. RESULTS

Results have been obtained at 300 K for a 100 Å well at different field strengths. Fig.1 and Fig.2 report the carrier drift velocity and the longitudinal diffusion coefficient as functions of field strength for both the 2D system and bulk GaAs. The absence of dissipative scattering mechanisms below the optical phonon temperature does not permit simulation of ohmic conditions. However the extrapolation to the low-field limit of the data for v_d and D satisfies the Einstein relation within the Monte Carlo accuracy.

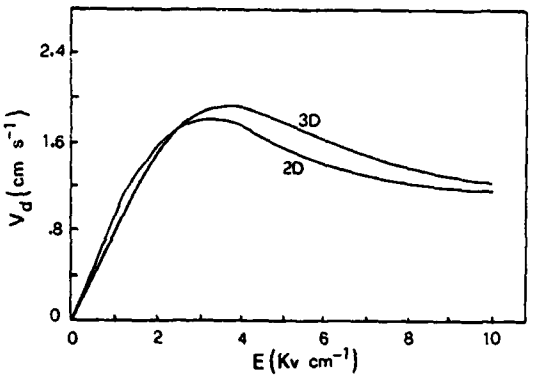


Fig.1: Drift velocity as a function of field strength at 300 K for the 2D and the 3D systems (see text).

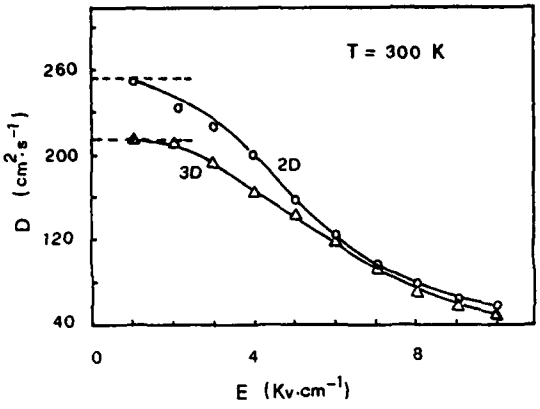


Fig.2: Longitudinal diffusion coefficient as a function of field strength at 300 K for the 2D and 3D cases. The orizontal lines represent the equilibrium values (see text).

The ohmic mobility in the two cases is not significantly different because the 2D polar optical scattering rate is very close to the same curve for 3D. Consequently the equilibrium D is also close between 2D and 3D.

A negative differential mobility effect is present in both curves, and the threshold field in 2D is lower than in 3D. At room temperature, for fields lower than the threshold for electron transfer to the upper valleys and bands the dominant scattering is essentially given by polar optical interaction. When the electron energy is high enough to allow transfer to upper valleys and upper subbands the mobility decreases with higher field strengths.

The diffusion coefficient is found to decrease monotonically in both cases at increasing field strengths. At low and intermediate fields D is larger in 2D than in 3D, because interband scattering introduces a new randomization process that increases diffusivity, noise, and decreases the autocorrelation time of velocity fluctuations. At higher fields the difference between 2D and 3D is reduced and finally it disappears when the electron energy is high enough to guarantee full randomization of electron paths in k-space.

Our results differ from the theoretical data of van Rheenen et al. [1] , who found large difference between 2D and 3D data for both drift velocity and diffusivity vs field, while we find better agreement with experimental data [2] , even though a direct comparison would require the knowledge of both experimental and theoretical low-field diffusivity.

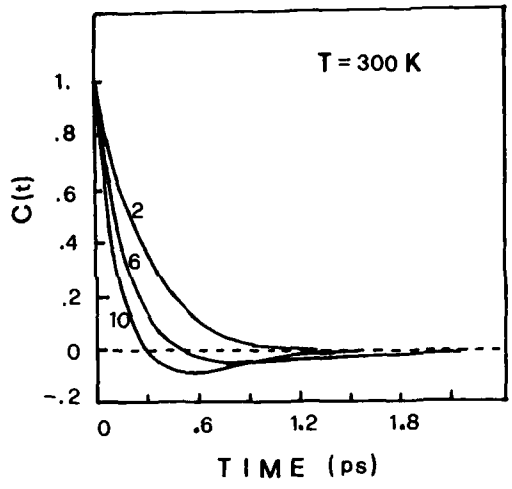


Fig.3: 2D autocorrelation function of velocity fluctuations as a function of time for E=2,6,10 kV/cm.

As a further confirmation of the above interpretation, Fig. 3 shows the normalised ACF of velocity fluctuations as a function of time for E=2,6,10 kV/cm for 2D and 3D A negative tail is present at 6 and 10 kV/cm which is due to the dynamical effect of the transfer back and forth from upper valleys. Furthermore the time necessary to cancel the correlations between velocity fluctuations is larger at fields close to the threshold field, and then it decreases at increasing fields.

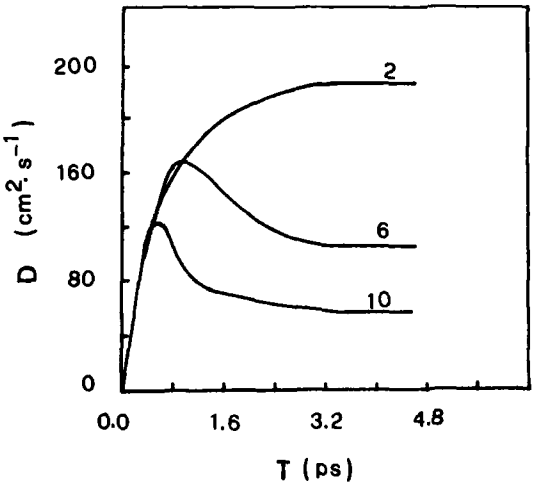


Fig.4: Transient diffusion coefficient as a function of time for the 2D system at the same fields shown in Fig.3.

Fig.4 shows the transient longitudinal diffusion coefficient as a function of time for the same fields of Fig.3. It is interesting to notice the presence of a bump at E=6 and 10 kV/cm related to the decrease in velocity during the transient, when the electron cloud begins the transfer to upper valleys and bands. In 3D GaAs D is found to be even negative because of the same effect, but in 2D the presence of upper subbands seems to smooth the strong negative correlation introduced by the usual dynamical picture, and D is always found positive.

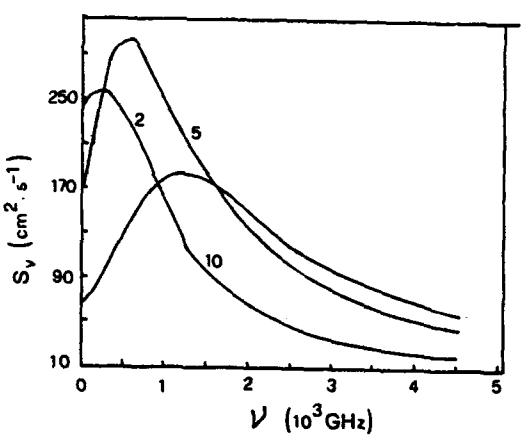


Fig.5: Power spectral density of velocity fluctuations as a function of frequenccy at E=2,5 and 10 kV/cm for the 2D case.

Finally Fig. 5 shows the power spectral density of velocity fluctuations as a function of frequency at E=2,5 and 10 kV/cm. This quantity exibits the same qualitative behaviour of the 3D case. The bump at low frequencies is due to the presence of the negative tail in the ACF.

ACKNOWLEDGEMENTS

We gratefully thank Carlo Jacoboni for helpful discussions and suggestions during the development of the research. The Computer Centers of The University of Modena and of the Oregon State University have provided computer facilities.

This work is partially supported by the European Research Office.

REFERENCES

1. A.D. van Rheenen and G. Bosman, Proc. Int. Conf. on Noise in Physical Systems and 1/f Noise, p.163, North Holland (1985).
2. C.Whiteside, G. Bosman and H. Morkoc, Proc. Int. Conf. on Noise in Physical Systems and 1/f Noise, Montreal, 1987 to be published.
3. F. Reif "Fundamentals of Statistical and Thermal Physics" Mc Graw Hill, New York, 1965.
4. R. Brunetti and C. Jacoboni, Phys. Rev. B29 (10),5739 (1984).
5. R. Brunetti and C. Jacoboni, in Semiconductors Probed by Ultrafast Laser Spectroscopy, vol.1, p.367 (1984).
6. C. Kittel "Elementary Statistical Physics", Wiley, New York (1958).
7. F.A. Riddoch and B. Ridley, J. Phys. C 16, 6971 (1983).
8. S. M. Goodnick and P. Lugli, Proc. Int. Conference on High Speed Electronics, Ed. Kallback and Beneking, Springer Series in Electronics and Photonics, Vol.22, 116 (1986).
9. S. M. Goodnick and P. Lugli, Phys. Rev. B, to be published.

A DETERMINISTIC PARTICLE METHOD FOR THE SEMICONDUCTOR BOLTZMANN EQUATION

Pierre DEGOND, Frederique GUYOT and Bernard NICLOT

Centre de mathematiques appliquees, Ecole Polytechnique, 91128 Palaiseau cedex, France.

We present a new particle method for the simulation of the semiconductor Boltzmann equation. It differs from the Monte-Carlo method by the approximation of the collision integral: with each particle, we associate a new degree of freedom, a weight, the time variations of which account for the collision integral by means of a quadrature formula. Linear as well as non-linear collision integrals can be handled the same way by this method. It has been applied to the polar optical scattering in the homogeneous field model. Numerical results are presented with an emphasis on the influence of the Pauli degeneracy terms.

1. INTRODUCTION

In modern submicronic devices, the physical assumptions which base the fluid models (e.g. the drift-diffusion or the hydrodynamic models) are no more fulfilled. Indeed the ratio of the mean free path of the carriers to the active length of the device is no longer small, and one has to consider the semiconductor Boltzmann equation, in order to achieve a physically reliable modelling. The most widely used numerical method for the simulation of the Boltzmann equation is the Monte-Carlo method (cf [1] and references therein), but iterative methods [2] and direct methods [3] have been successfully used in very particular cases.

The Monte-Carlo method, although very successful in many areas, presents two major drawbacks: first, the stochastic nature of the method introduces a lot of numerical noise which makes the computations very uneasy in the transient inhomogeneous regimes; second, the definition of an appropriate stochastic process can be extremely complicated in the non-linear cases and lead to quite expensive computations. Thus the development of new direct numerical methods seems very important for the future of the numerical modelling of the Boltzmann equation.

This paper intends to present a new method which could provide an interesting alternative to the Monte-Carlo method (cf also [4]). It is based on the ideas of Raviart and Mas-Gallic [5], and is presented in the next section .

2. PRESENTATION OF THE NUMERICAL METHOD

In this paper, we will only consider a constant field model. This model describes an infinite and homogeneous sample of semiconductor submitted to a constant external electric field. However the method extends quite straightforwardly to inhomogeneous cases [5], and numerical experiments are currently performed in this context. We recall that the Boltzmann equation in the constant field geometry is written:

$$(1)\quad \partial_t f(k_1, k_2, t) - qE/\hbar\, \partial_{k_1} f(k_1, k_2, t) = Q(f)$$

where the distribution function f depends on the components k_1 and k_2 of the electron wave vector k, respectively parallel and tranverse to the electric field E. We recall that q is the (positive) electronic charge and $\hbar$ is the reduced Planck constant.

The collision term Q(f) describes the interactions between the electrons and the lattice (up to now, we do not consider electron- electron collisions), and is given by:

$$(2)\quad Q(f) = \int [\, s(k',k)\, f(k')\, (1-f(k)) - s(k,k')\, f(k)\, (1-f(k'))\,]\, dk'$$

where $s(k,k')$ is the transition probability from a state k to a state k'. The (1-f) terms in (2) account for Pauli's exclusion principle which also implies that f is comprised between 0 and 1. For lightly doped materials, the (1-f) terms can be neglected and the collision operator approximated by:

$$(3)\quad Q(f) = \int [\, s(k',k)\, f(k') - s(k,k')\, f(k)\,]\, dk'$$

The numerical approximation relies on a particle description of the distribution function, which means that f is decomposed into a sum of delta functions:

$$(4)\quad f \simeq \Sigma_i\ \omega_i\ f_i(t)\ \delta(k_1 - k_1^i(t))\ \delta(k_2 - k_2^i(t))$$

where $k_1^i(t)$ and $k_2^i(t)$ are the positions (in the k space) of the i-th particle, ω_i, a constant control volume, and $f_i(t)$ a time dependent weight. In a Monte-Carlo method, the product $\omega_i f_i$ is constant in time, and both the collisionless dynamics (given by the left hand side of (1)), and the collisions (right hand side) are accounted for by the motion of the particles. In the present method, the motion of the particles is ruled by a collisionless dynamics, whereas the collisions are accounted for by the variation of the weights, through a quadrature evaluation of the integral (2):

$$(5)\quad \partial_t k_1^i = -\, qE/\hbar \qquad ; \qquad \partial_t\, k_2^i = 0$$

$$(6)\quad \partial_t f_i = \Sigma_j\ [\, s^\alpha(k_j,k_i)\, f_j\, (1 - f_i) - s^\alpha(k_i,k_j)\, f_i\, (1 - f_j)\,]$$

$s^\alpha(k,k')$ represents a smoothed approximation of $s(k,k')$ when it displays singularities (cf details in [4]). The differential system (5) (6), is solved by standard finite difference methods. This method presents strong analogies with that of Kuivalainen and Lindberg [6].

3. NUMERICAL RESULTS FOR POLAR OPTICAL SCATTERING

The aim of this section is to present an application of our numerical method for the investigation of degeneracy effects on polar optical scattering. We used the expression of the transition rate $s(k,k')$ given in [1] and the numerical values for GaAs [1]. Beside the polar optical scattering term a relaxation time model (of order 10^{-11} s) was introduced to account for the other sources of scattering (impurities, acoustical phonons). The distribution function was initiated with an equilibrium Fermi-Dirac distribution at the lattice temperature (T = 77 K) and the electric field was 10^5 Vm^{-1} which is the highest value for which we can neglect intervalley effects.

We performed several numerical computations with different electron densities (that is different values of the chemical potential for the initial Fermi-Dirac distribution). On figure 1, we display the mean velocity versus time for 3 values of the electron density. At a density of 10^{17} cm^{-3} (plot 1), a slight velocity overshoot is present. At 10^{18} cm^{-3} (plot 2), the stationary velocity is smaller and the overshoot has vanished. But the qualitative behaviour is quite the same. However, at 3.10^{18} cm^{-3} (plot 3) the time of the simulation was not sufficient to reach the stationary state. Furthermore, the stationary velocity is much higher than in case 2.

The physical interpretation of this behaviour is given when considering figures 2 and 3 which display the traces of the distribution function along the k_1 axis (parallel to the electric field) respectively at the initial time and at time t = 3 ps and for the 3 values of the density. On the figures, we have reported the value of the threshold for polar optical emission (ie the wave-vector corresponding to the optical phonon

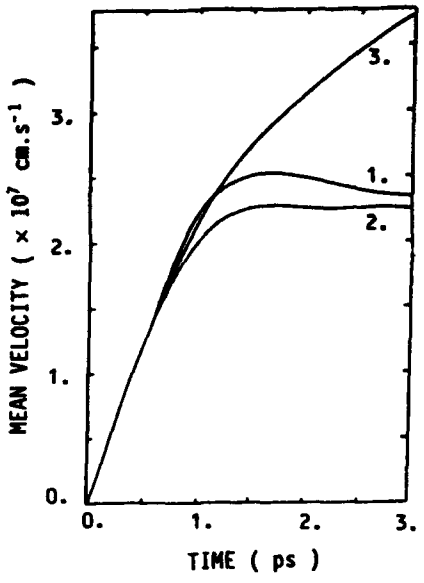


Fig 1: mean velocity versus time: density = 10^{17} cm^{-3} (plot 1) , 10^{18} cm^{-3} (plot 2) , 3.10^{18} cm^{-3} (plot 3).

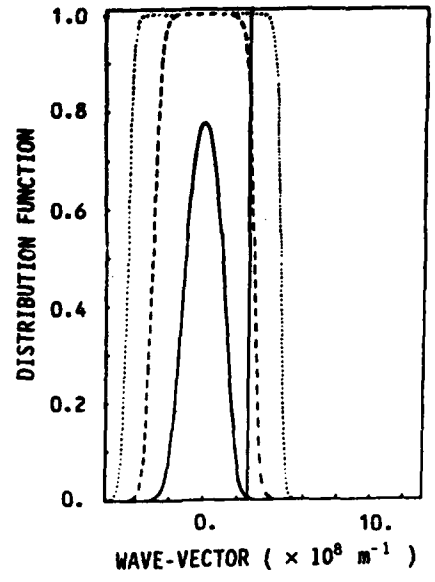


Fig 2: trace of the distribution function along the k_1 axis at time t=0: density = 10^{17} cm^{-3} (solid line) , 10^{18} cm^{-3} (dashed line) , 3.10^{18} cm^{-3} (dotted line). The vertical line indicates the position of the threshold for polar optical phonon emission.

energy: vertical line). Indeed, the scattering becomes significant only for electrons with wave-vectors larger than the threshold.

For a low density (case 1), the chemical potential is much lower than the threshold energy (figure 2). Thus, the transient regime begins with a quasi ballistic part, where electrons are accelerated by the field with almost no scattering, until they reach the threshold energy. Then, they undergo a strong scattering which sends them back to lower energies and produces the overshoot on figure 1. The stationary distribution function (figure 3) has a very sharp gradient near the threshold, indication of its very strong influence. In this case, the use of a linear scattering term (3) would produce very similar results [4].

At the intermediate density (case 2), the chemical potential is almost equal to the threshold energy (figure 2), which lies at the edge between full and empty states. Thus, a

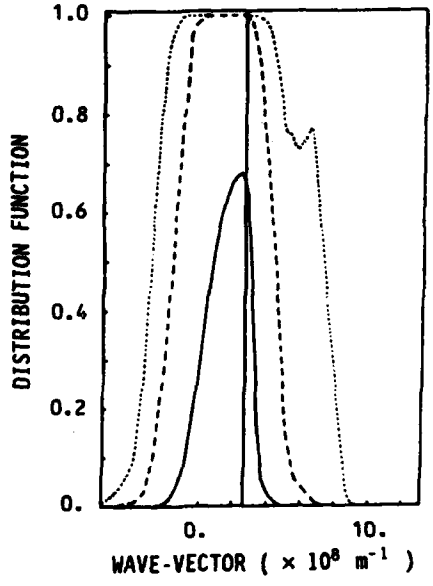


Fig 3: same as fig 2 at time 3 ps.

larger number of electrons undergo optical scattering, from the outset. This reduces the ballistic phase of the motion and the overshoot disappears. Similarly, the stationary distribution (figure 3) does not keep any reminiscence of the optical threshold.

At the highest density (case 3), the chemical potential is much higher than the threshold energy (figure 2) the importance of which has completely vanished: the qualitative behaviour of the velocity curve (figure 1), is then similar to that of a relaxation time model. On the other hand the (1-f) terms in expression (2) reduce strongly the magnitude of the interaction, resulting in a dramatic increase of the relaxation time and of the stationary velocity. The distribution function at time 3 ps is displayed, but its peaked structure is not understood.

In the last case, it is believable that the results would be changed by the incorporation of a more real model of impurity scattering. However, it seems that this would not affect our qualitative conclusions and furthermore, this was essentially intended to give an example of the capabilities of the method in the non-linear cases. An intervalley scattering case has also been investigated and is detailed in [4]. The electron-electron interaction is currently investigated.

4. CONCLUSION

We have proposed a new numerical method, which is complementary to the Monte-Carlo method in that it seems to be more straightforwardly extended to non-linear collision operators. Furthermore, this method enables to obtain snapshots of the distribution function during its time evolution, which could be of great help for the understanding of physical phenomena. As an example, we have applied it to the analysis of polar optical scattering under high doping conditions and found that these could largely affect the behaviour of the interaction.

ACKNOWLEDGEMENTS

The computer cost has been supported by the "Centre de calcul vectoriel pour la recherche", France.

REFERENCES

[1] REGGIANI, L. (ed), Hot electron transport in semiconductors (topics in applied physics series, Springer, Berlin, 1985).

[2] REES, D.H., J. Phys. Chem. Solids 30 (1969), 643.

[3] BARANGER, H.U. and WILKINS, J.W., Phys. Rev. B 30 (1984), 7349; Physica 134 B (1985), 470; Appl. Phys. Lett. 49 (1986), 176.

[4] NICLOT, B., DEGOND, P., POUPAUD, F., Deterministic particle simulations of the Boltzmann transport equation of semiconductors, preprint.

[5] MAS-GALLIC, S., A deterministic particle treatment of the linearized Boltzmann equation, preprint, to appear in Tranp. Th. Stat. Phys.

[6] KUIVALAINEN, P. and LINDBERG, K., in Kullback, B. and Beneking, H. (eds.), High speed electronics (Springer Verlag, New-York, 1986).

Session P1.1

Posters

Monday, September 14, 1987

TiW as a direct contact material and a diffusion barrier to n+ and p+ implanted areas

A.Lindberg, M. Östling, H. Norström* and U. Wennström*

Institute of Microwave Technology,
Box 70033, S-100 44 Stockholm, Sweden

* RIFA AB, S-163 81 Stockholm, Sweden

Sputtered TiW(15/85 wt.%) has been investigated as a diffusion barrier layer between Al and Si and as a direct contact material to n^+-and p^+-doped silicon areas. Rutherford backscattering spectroscopy (RBS) in combination with reverse leakage current measurements on gated diodes were used to investigate the barrier properties of the deposited TiW films. TiW was observed to prevent intermixing between Al and Si up to 450 °C. Four terminal Kelvin resistor structures, implanted with different doses of As and BF_2, ranging from 1E15 cm^{-2} up to 1E16 cm^{-2}, were used to determine the respective contact resistivity. The actual surface doping concentration was determined by spreading resistance profiling (SRP). Contact resistance values well within the acceptable limit for VLSI processing were achieved.

1. INTRODUCTION

To prevent contact failure of shallow junctions as a result of spiking or formation of silicon precipitates, a diffusion barrier film, positioned between the aluminium interconnect layer and the silicon contact, must be added. Among the barrier systems investigated TiW [1] and TiN [2] have achieved considerable interest.

TiW has been extensively used in bipolar technology to prevent intermixing between the Al-interconnect layer and PtSi Schottky-clamped contacts. Recent reports [3,4] have demonstrated the possible use of TiW both as a diffusion barrier layer and as a contact material to n^+- and p^+-diffusions in advanced CMOS processing. However, only limited amount of information is presently available on the contact resistivity of TiW. Furthermore, the published data exhibits a large spread, possibly due to the various methods used to define the contact windows [5].

Experiments were therefore performed to study the influence of surface dopant concentration on the contact resistivity of TiW. A parallel investigation of the diffusion barrier properties of TiW was also carried out.

2. EXPERMENTAL DETAILS

The experiments in this study were performed on both p- and n-type silicon wafers, <100> oriented of 15-25 Ωcm resistivity. The active device area were patterned using conventional LOCOS technique. The n-type wafers were implanted with BF_2 and the p-type wafers were implanted with arsenic. The implantation energy was 50 KeV for both arsenic and BF_2. The doses used were 1E15, 2E15, 5E15 and 1E16 cm^{-2}. Ion implantation was performed through a thin (200Å) screen oxide. The arsenic implantations were annealed at 1000 °C, 30 minutes, in N2-ambient and the BF_2 implantations were nitrogen annealed at 925 °C, 40 minutes.

A phosphosilicate glass (PSG) film, approximately 4500Å, was subsequently deposited on the wafers, to serve as a passivation.

Contact windows were opened by reactive ion etching (RIE) in a mixture of CHF_3-O_2.

Before metal deposition the wafers were subjected to a standard cleaning proceedure followed by a short dip in diluted HF. Three different sets of samples were prepared. The first group consisted of 3500Å of Al covered with a top layer of 1100Å of $MoSi_x$. The composition of the $MoSi_x$ film was determined by RBS to be $MoSi_{1.9}$. The $MoSi_{1.9}$ top barrier was include to serve as an etch stop and hillock supressor in a double layer metallization concept.

In the second group, approximately 1200Å of TiW was inserted between the Al/$MoSi_{1.9}$ sandwich and the silicon to act as a barrier layer. The trilayer structure was deposited without breaking vacuum.

The third group was similar to the second one except that after TiW deposition the wafers were taken out of the sputtering system and submitted to an oxygen plasma treatment - TiW(O_2), at 1 Torr, 450 W for 30 minutes, after which they were returned to the sputtering system for the Al/$MoSi_{1.9}$ sandwich deposition. The plasma treatment was done in order to improve the barrier properties of TiW.

A D.C. magnetron sputtering system was used for metal deposition. The substrate temperature during sputtering was kept at 300 °C. The Al films were deposited from a Al(1% Si, 0.5% Cu) target. The $MoSi_{1.9}$ and TiW films were deposited from compound targets ($MoSi_{2.1}$ and TiW 15/85 wt. %). All samples were given a thermal sintering treatment at 400 °C for 75 minutes in hydrogen prior to electrical evaluation.

Kelvin structures were used to measure the contact resistance and gated diodes (500x500 µm) were used to measure the reverse leakage current.

Part of the samples were isochronically annealed for 60 minutes at temperatures up to 550 °C in order to investigate the thermal stability of the respective metal systems.

3. Results

3.1 Compositional analysis

The deposition parameters of the TiW barrier layer was investigated by means of Rutherford backscattering spectroscopy (RBS). This analysis was performed on a Tandem van de Graaf accelerator operated at 2.4 MeV, $^4He^+$-ions. The ion current was in the 20-40 nA range. Figure 1 shows the titanium content in the deposited TiW film verus pressure and power. It is readily seen that the tungsten content in as deposited films increases when the pressure is reduced as a result of less gas phase scattering [6].

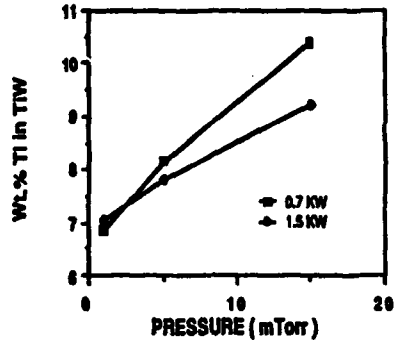


FIGURE 1

The weight % of Ti in the deposited TiW film vs. argon pressure and power.

3.2 Thermal stability

The thermal stability of the barrier structure was also analyzed by RBS. The backscattering results of the Al/TiW/silicon structure is shown in figure 2a. Three spectra are displayed, the as deposited structure, after annealing at 450°C and 475°C for 1h respectively. The RBS spectra for the as-deposited and the 450°C samples are almost identical, indicating no interreactions at all. Annealing at 475°C leads to a reaction between Al and TiW, as can be seen from the push forward of the leading edge for the W signal and from the tail of the trailing edge for the Al signal. In one set of samples the TiW film was subjected to an O_2-plasma treatment

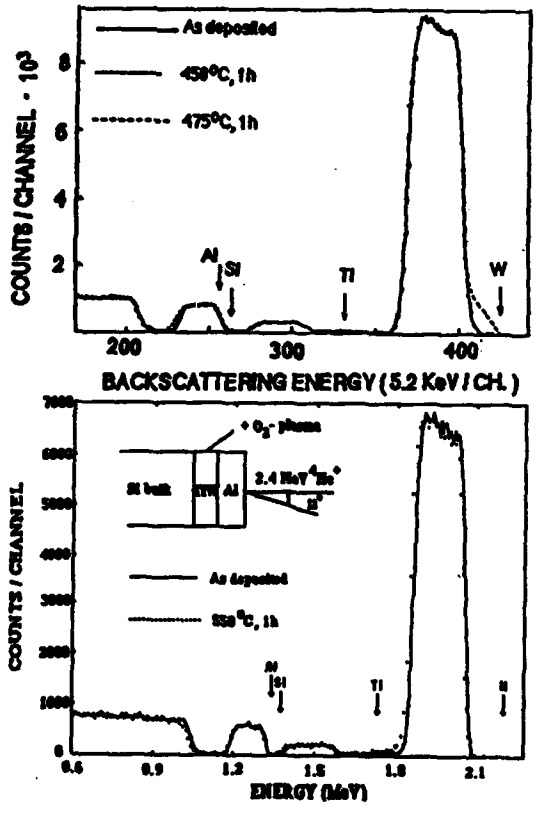


FIGURE 2

Rutherford backscattering spectra showing the metallization system before and after heattreatments. The upper figure displays the results for the Al/TiW/silicon system and the lower figure shows the Al/TiW(O_2)/silicon system.

before the aluminium deposition. Figure 2b shows the RBS results. It is readily seen from these spectra that such a barrier layer can withstand significantly higher annealing temperatures viz.up to 550°C. Only a minor tail in the tungsten signal can be seen after heattreatment at 550°C, indicating a small interreaction.

Backscattering results, however, only comprise information from unpatterned structures. Diodes were therfore prepared in order to electrically evaluate the barrier properties of the metallization scheme. The results from reverse leakage current measurements are presented in figure 3 where the relative fraction of failing diodes are plotted versus annealing temperature. A diode is considered as failed if the reverse current density exceeds 0.2 μAcm^{-2}.

Three structures are shown: i) Al/Si, ii) Al/TiW/Si and iii) Al/TiW(O_2)/Si. All structures have a top metallization layer of $MoSi_{1.9}$. For the Al/TiW/Si system the diode failure population starts to raise at annealing temperatures above 450 °C. When the TiW layer was subjected to the O_2-plasma treatment a very low failure rate was detected even for temperatures up to 550 °C. These observations concur well with the RBS measurements (fig.2).

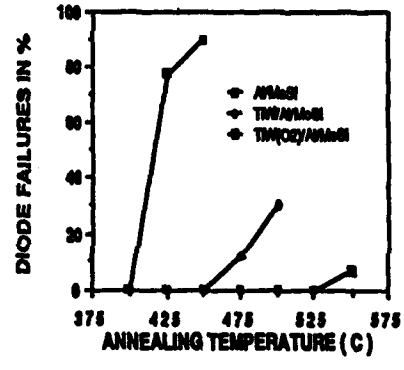


FIGURE 3

Failure % for n^+-diodes (5E15 cm^{-2}) versus annealing temperature.

3.3 Contact properties

The four terminal Kelvin resistor structure [7] was used to extract the metal-semiconductor contact resistance R_c. A DC- parameter analyzer was used to force a known current through the contact on two opposite pads and simultaneously record the resultant voltage drop between the remaining pads orthogonal to the current path. The specific contact resistivity (ρ_c) was obtained by multiplying the measured contact resistance R_c, with the known contact area, A, i.e.

$$\rho_c = R_c \cdot A \qquad (1)$$

The contact resistance as a function of contact hole size, for a fixed dose of 5E15 cm^{-2}, are plotted in figure 4a and 4b. As expected, oxygen plasma treated TiW contacts exhibites the largest resistance. It is also noted that aluminium, as a contact material to p^+ silicon surpasses ordinary TiW whereas the reverse holds for contacts made to n^+ doped areas.

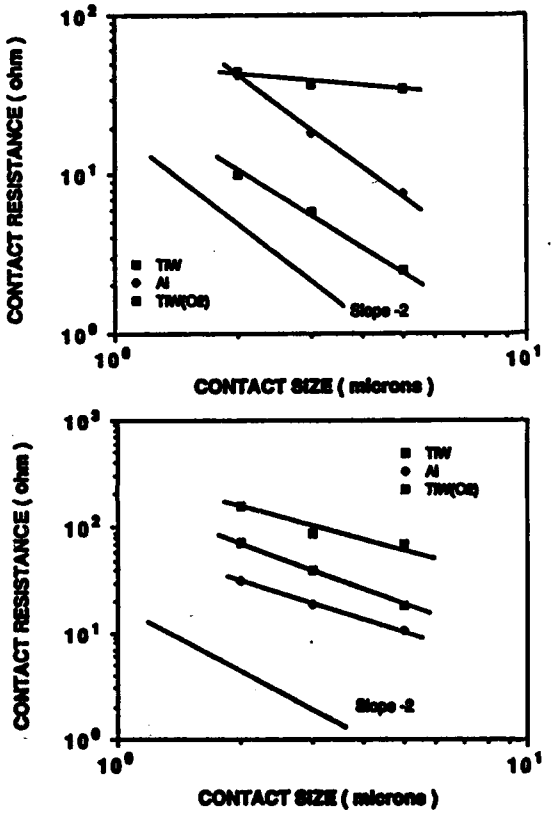


FIGURE 4

Contact resistance as a function of contact hole size. The figure shows the results for arsenic (upper fig.) and BF_2 (lower fig.) implantations.

The resistance is observed to increase less dramatically than predicted by the scaling theory in agreement with the findings of Cohen et al. [8]. The deviation from a strictly area dependent contact resistance value is largest for Al-contacts to p^+ silicon. This could possibly be the result of an area increase due to aluminium pitting [9] upon anneal as the TiW contact resistance reveals a larger dependence on contact size.

To investigate the influence of dopant concentration on the contact resistance, Kelvin patterns were implanted with various doses ranging from 1E15 cm^{-2} up to 1E16 cm^{-2}. The active surface concentration, subsequent to thermal activation, was estimated by spreading resistance profiling (SRP). The results are recapitulated in table 1 together with SUPREM simulation data.

	MEASURED			SUPREM		
Dose (cm^{-2})	Sheet resist. (Ω/sq.)	Surf. conc. (cm^{-3})		Sheet resist. (Ω/sq.)	Surf. conc. (cm^{-3})	Type
1E15	170	4.0E19		89	8.0E19	As
2E15	79	6.5E19		50	1.3E20	As
5E15	38	8.5E19		26	2.1E20	As
1E16	26	1.2E20		20	3.1E20	As
1E15	270	1.6E19		195	1.0E19	BF_2
2E15	143	2.2E19		99	2.1E19	BF_2
5E15	93	3.7E19		39	5.4E19	BF_2
1E16	82	4.5E19		21	1.0E20	BF_2

TABLE 1

Sheet resistivity and surface concentration verus dose. Measured and simulated using SUPREM.

Based on tunnling theory [10,11], the specific contact resistivity ρ_c can be expressed as

$$\rho_c = A \cdot \exp(B \cdot \Phi_b / \sqrt{C_o}) \qquad (2)$$

where A and B are different constants, Φ_b denotes the metal-semiconductor barrier height and C_o denotes the active surface concentration.

In figure 5a and 5b the recorded data of $\log(\rho_c)$ is plotted versus $1/\sqrt{C_o}$ for two different contact sizes. The presented values closely resembles a straight line as predicted by the equation (2) above.

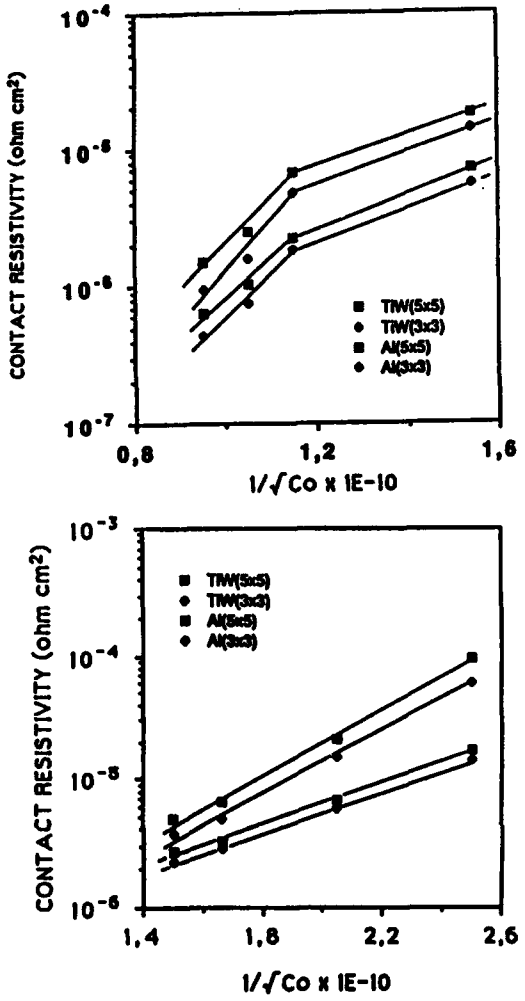


FIGURE 5

The contact resistivity is plotted vs. $1/\sqrt{C_o}$ for different contact areas. The upper diagram shows the results for arsenic implantation whereas the lower diagram shows the results for the BF_2 implantations.

4. CONCLUDING REMARKS

For a given contact resistivity the ratio of the contact resistance to the intrinsic resistance of an FET increases quadratically if constant field scaling is applied [12].

In order to keep the contribution from the contact resistances below 1% of the channel resistance for a 1 µm technology, Chen and Rensch [12] have shown that a contact resistivity to n+ silicon below 8E-7 would be necessary.

A slightly higher value of the contact resistivity to p+ areas is acceptable because of the lower transconductance of p channel devices.

The presented resistivity data clearly indicates that TiW can be used as a direct contact material to n+ silicon at and above the one micron level where the resistivity of aluminium contacts would be too high. However it might be necessary to improve the contact between p+ silicon and TiW, e.g. use rapid thermal processing for activation of boron in combination with an improved contact etch proceedure [5].

REFERENCES

[1]. P.B.Ghate, J.C. Blair, C.R. Fuller and G.E. McGuire, Thin Solid Films, 53, (1978) pp.117

[2]. M-A. Nicolet, Thin Solid Films, 52, (1978) pp.415

[3]. S.S. Cohen, M.J. Kim, B. Gorowitz, R. Saia and T.F.McNelly,Appl.Phys.Lett.,45.4, (1984) pp.414

[4]. R.A.M. Wolters and A.J.M. Nellissen, Solid State Technology, 29, 2, (1986) pp.131

[5]. M.J. Kim, D.M. Brown, S.S. Cohen, P. Piacenta and B. Gorowitz, IEEE Trans. Electron Dev., ED-32, (1985) pp.1328

[6]. M. Hill, Solid State Technology, pp.53 (1980)

[7]. S.J. Proctor and L.W. Linholm, Electron Dev. Lett., EDL-3, (1982) pp.294

[8]. S.S. Cohen, G. Gildenblat and D.M Brown, J. Electrochem.Soc.,130, 4, (1983) pp.978

[9]. S.J. Proctor, L.W. Linholm and J.A. Mazes, Trans.El. Dev., ED-30, (1983) pp.1535

[10]. A.Y.C.Yu, Solid State Electron., 13, (1970)pp.239

[11].C.Y. Chang, Y.K. Fang and S.M. Sze, Solid State Electron., 14, (1971) pp.541

[12].J. Y-T. Chen and D.B. Rensch, IEEE Trans. Electron Dev., ED-30, 11. 1983, pp.1542

P1.1.2

COMPARISON OF THE BEHAVIOUR OF ARSENIC DURING TITANIUM AND TUNGSTEN DISILICIDE FORMATION

J. TORRES, J.C. OBERLIN, G. BOMCHIL, A. PERIO
D. LEVY*
A. SAULNIER, J.P. PONPON, R. STUCK**

Centre National d'Etudes des Télecommunications BP 98, 38243 MEYLAN FRANCE
Bull, Rue Jean Jaurès, 78340 LES CLAYES-SOUS-BOIS FRANCE*
Groupe Phase, u.a. CNRS 292, Centre de Recherches Nucléaires, 67037 STRASBOURG**

The direct reaction between W or Ti and silicon, to form tungsten silicide and titanium silicide is studied for both unimplanted and arsenic implanted silicon substrates. For W on implanted substrates, the silicide growth rate decreases drastically, being dependent on the implantation dose and energy. The dopant effect can be explained by the formation of a dopant rich phase at the silicide/silicon interface. For Ti, the dopant effect is much less pronounced, slowing down the rate of formation of the silicon rich phase $TiSi_2$. A clear arsenic pile up at the silicide-silicon interface is not observed. The diffusion barrier is distributed in the whole silicide layer.

INTRODUCTION

The direct reaction between refractory metals and silicon to form a silicide is becoming an important step in the preparation of self aligned contacts for VLSI technology. As the silicidation is generally realized over highly doped silicon regions (source and drain in MOS transistors), it is necessary to precisely control the behaviour and the role of the impurities during the heat treatments. Specially, the influence of dopants on the silicidation mechanism and on the reaction kinetics is of prime interest as they can modify the silicon diffusion through the already formed silicide layer and strongly delay or even block the reaction between the metal and the silicon. In addition, redistribution effects may lead to dopant losses and changes in the junctions properties.

As the behaviour of a refractory metal-silicon system will depend on the nature of the metal, we have investigated in this paper the growth of titanium and tungsten disilicides in presence of implanted arsenic in order to have a comparison of these two structures in terms of silicidation kinetics and arsenic redistribution. Due to their quite different behaviour with respect to oxygen, titanium and tungsten represent two extrem cases for the reaction of a refractory metal with silicon. Nevertheless they are both promising candidates for salicide applications so that a comprehensive knowledge of the nature of their reaction with undoped and highly doped silicon will be necessary to achieve reproducible and controled properties (silicide phase obtained, thickness, sheet resistance...) during the formation of contacts in a process at the micrometer scale.

The W/Si reaction

A reference sample was prepared by sputtering 1000 A° of tungsten on an unimplanted monocrystalline silicon substrate. For a first group of samples, the silicon substrate has been arsenic implanted at 180 kev using three doses prior to a 400 A° W layer deposition (table I). The implantation energy was choosen to localise the arsenic peak at the final silicide/silicon interface position when complete WSi_2 formation is achieved. For a second group of samples, a 400 A° thick W layer was deposited on silicon substrates implanted at thee different energies, the implantation depths decreasing from the preceeding value down to a very shallow implantation. The samples were furnace annealed a 99,99% pure Argon flow at

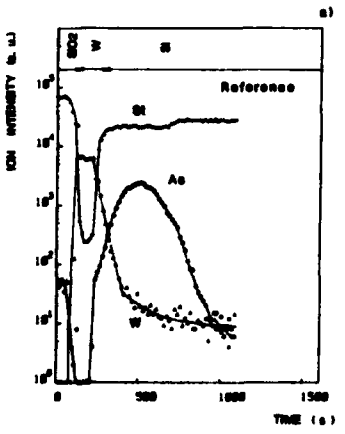
temperatures ranging under from 600°C up to 1000°C.

DOSE ATOMS	DEPTH nm	T °C	TIME H	THICKNESS nm
3E15	100	725	1	< 20
3E15	100	725	22	40
3E15	100	900	1	70
9E15	100	900	1	30
3E16	100	900	1	NO REACTION
6E15	37	850	2	NO REACTION
6E15	70	850	2	< 20
6E15	110	850	2	100

TABLE I

Implantation conditions and annealing results in the W/Si system.

For the reference sample, the disilicide formation is already observed at 600°C and after one hour annealing at 675°C, the reaction is nearly complete. The temperature and time needed to obtain a complete reaction between W and arsenic implanted silicon substrates are considerably greater. For the group I samples, at a annealing temperature of 900°C, the amount of silicide formed for one hour decreases with increasing As doses. On the other hand for the group II samples the silicide amount formed after annealing at 850°C for two hours decreases when decreasing the implantation depth. RBS and SIMS analysis performed on group II samples (Fig I) indicate a high level of As segregation into the Si substrate with a clear dopant accumulation at the growing silicide/silicon interface (snow-plow effect).



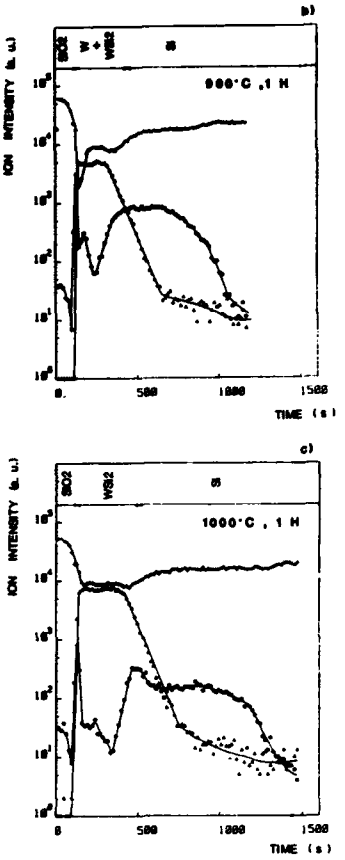


FIGURE 1

SIMS profiles of SiO_2/W/5E15 ion/cm^2 arsenic doped Si substrate structures (Rp≃79nm)
a) unannealed ; b) annealing at 900°C for one hour ; c) annealing at 1000°C for one hour.

The greater is the arsenic dose initially located in the silicon consumed during the silicide formation, the greater is the arsenic accumulated at the silicide/silicon interface. The X-ray diffraction spectra of the sample implanted with the highest dose show additional peaks corresponding to the W_2 As_3 phase.

Ti/Si reaction

Arsenic doses of 2E 15 ion/cm^2 and 1 E 16 ion/cm^2 have been implanted into silicon prior to titanium deposition. Titanium layers up to 100nm thick have been e-beam evaporated under vacuum. In order to reduce oxygen incorporation during the deposition process, the substrates were held at 350°C. Annealing was then performed

at 700°C and 750°C for 1 to 50 seconds in a tight rapid thermal annealer (HEATPULSE 410) under argon ambient to avoid any titanium contamination by oxygen during the heat treatment. The growth kinetics of the silicide and the arsenic redistribution have been investigated by RBS and SIMS as a function of the annealing conditions.

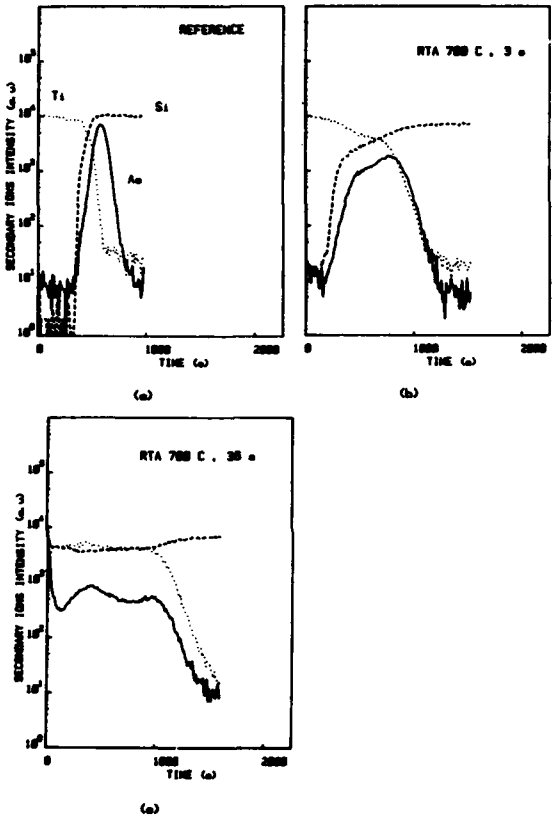


FIGURE 2

SIMS profiles recorded for 2 E15/cm^2 arsenic atoms at 40 keV introduced into silicon before Ti deposition.
a) unannealed ; b) annealing at 700°C for 3 seconds ; c) annealing at 700°C for 36 seconds.

As shown in Fig 2, for a shallow implantation into silicon, arsenic incorporated in the growing silicide, redistributes very quickly through this layer. At the end of the reaction, the arsenic is found uniformly distributed into the silicide, this result being independent of the initial position of the arsenic implantation provided it is shallower than the final silicide/silicon interface position. At the same time the silicidation kinetics is somewhat slowed down (Fig 3). This effect increases with the arsenic dose and decreases with the annealing temperature. However, in contrast to what can be observed for the W-Si system, the reaction between Ti and Si is not hindered but the rate of formation of the silicon rich phase $TiSi_2$, is noticeably reduced, especially for the highest arsenic dose. As a result, the time needed to completely transform the titanium into $TiSi_2$ for the highest implanted substrates, is actually much larger than for implanted samples. Fig 3 shows that at 700°C, this time increases from about 12 to 40 seconds for a 2E15 ion/cm^2 dose. Furthermore for the 1E16 ion/cm^2 dose, the reaction is not complete within such an annealing duration and both $TiSi_2$ and a metal rich silicide, probably TiSi still remain. Obviously, as the lowest sheet resistance values can only be obtained with the disilicide phase, the presence of large concentrations of arsenic delays the achievement of convenient electrical properties, although they do not actually impede the reaction.

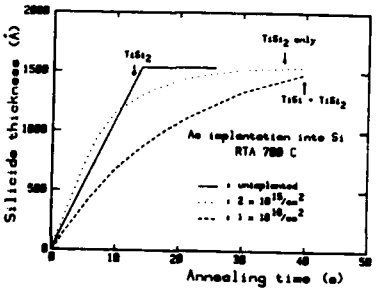


FIGURE 3

Kinetics of silicidation by Ti/Si reaction as a function of arsenic dose implanted in Si. Arrows indicate the final silicide phases observed.

Discussion

As shown by the experimental results, arsenic implantation in the silicon substrate drastically influences the W and Ti silicide formation, slowing down the silicide growth kinetics or even blocking it. During W silicide growing, as the silicon atoms leave the

silicon/silicide interface to diffuse through the formed silicide layer and react with metal atoms at the metal/silicide interface, the impurity is left behind and accumulates at the silicide/silicon interface. Above a critical arsenic level accumulated at this interface, the formation of the W_2As_3 arsenide compound can take place. Silicon diffusion is blocked thus also silicide formation. The silicide growth can be stopped uniformly leaving an homogeneous unreated W film, or the reaction can proceed only at a few points through the diffusion barrier, likely W_2As_3, built up at the silicide/silicon interface leading to a laterally uniform metal layer [1]. In an intent to overcome the barrier effect, samples of group II and III were annealed at temperatures higher than 900°C. In this case, large stresses develop probably at the W/WSi_2 boundaries [2] leading, in several cases, to non adherent layers.

The dopant segregation clearly observed in the W/Si reaction which produces an As pile-up acting as a diffusion barrier at the silicide/silicon interface, is believed to occur also for the Ti/Si case [3]. R.B.S observations of such arsenic segregation between $TiSi_2$ and Si have been recently reported [4]. However, the SIMS results we have obtained do not show that such accumulation is very important. In addition, as the dopant effects are much less pronounced for Ti/Si than for W/Si, it can be supposed that the process leading to a reduction of the $TiSi_2$ growth kinetics may be different from the W/Si case. Then, in agreement with recent results [5], it seems more reasonable to assume that the delay in titanium silicide growth is related to the presence of the impurity in the silicide layer rather than at the interface. An arsenic-titanium compound distributed in the silicide, at the grain boundaries, will not stop the reaction but only slows it by blocking easy diffusion paths and reducing silicon transport toward the unreacted titanium. To check this assumption, arsenic has been implanted into Ti instead of into Si, and the samples have been annealed at 700°C and 750°C for 1 to 50 seconds. The silicide growth kinetics results are similar to those obtained with As implanted in silicon. In contrast, when As is implanted into W, the reaction between W and Si occurs without any blockage and the silicide growth kinetics is quite similar to those obtained with the non implanted reference sample.

Conclusion

It has been shown that arsenic has a great influence on the reaction between W and Ti with silicon to form the disilicide. In the case of W, the silicide growth is drastically slown down or even blocked. Similar trends have been found for Ti. It seems however that the diffusion barrier which prevents the silicide formation is less effective. The results have been explained by assuming that in the Ti case, the diffusion barrier is delocalised and distributed in the whole silicide layer whereas in the W case, a discrete dopant rich blocking layer is formed at the silicon/silicide interface.

References

[1] C.D. Lien and M.A. Nicolet, J.Vac.Sci.Technol.B2, 738 (1984)

[2] G. Bomchil, G. Goltz and J. Torres, Thin Solid Films 140, 59 (1986)

[3] H.K Park, J. Sachitano, M. Mc Pherson, Y. Yamaguchi, G. Lehman, J. Vac. Sci. Technol. A2, 264 (1984)

[4] K.Maex, L. Van den Hove, R.F. De Keersmaecker, J.C. Oberlin, A. Perio, J. Torres, G. Bomchil, W.F. Van der Werg, Special issue "Le vide, les couches minces", Europeen Workshop on refractory metals and silicides, Aussois, 1987, p.95

[5] R. Beyers, D. Coulman, P. Merchant, J. Appl. Phys. 61, 5110 (1987)

LPCVD TUNGSTEN FILLED VIAS FOR MULTILAYER METALLIZATION

S.-L. ZHANG*, R. BUCHTA*, T. JOHANSSON*, H. NORSTRÖM**, and U. WENNSTRÖM**

* Institute of Microwave Technology, S-100 44 Stockholm, Sweden
**RIFA AB, S-163 81 Stockholm, Sweden

ABSTRACT

Tungsten, deposited in a hot wall LPCVD furnace, was used to fill vias and contact holes. Starting from a planarized silicondioxide surface, vias and contact holes were opened by RIE to the first metallayer. Prior to tungsten deposition, a layer of 500Å TiW was sputter-deposited onto the wafers. The thickness of LPCVD tungsten was chosen to result in a planar surface with all the vias filled. In the next step tungsten and TiW were backetched, using a mixture of CF_4/O_2, thus reestablishing the planarized silicondioxide surface with all the vias and contacts filled with tungsten.Electrical measurements comfirmed good, uniform contacts with low specific contact resistance. The results were successfully implemented in a commercially available CMOS process.

INTRODUCTION

Several methods have been proposed to overcome the metal step coverage problem in multilayer metallization systems[1,2,3]. For vertically etched vias and contacts of varying depth, CVD processes are superior to other deposition methods because they give conformal coating and allow batch processing. Recently, interest has been focussed on CVD tungsten[3], which despite of its relatively high resistivity, finds an increasing number of applications[4].

In this paper we describe a hot wall CVD tungsten process to fill via holes and contact windows, which in combination with a backetch procedure leads to an isoplanar surface, thus facilitating the subsequent aluminium metallization.

EXPERIMENTAL

Vias and contact holes with a diameter of 1.5 µm were opened with RIE on a planarized wafer down to the first metal layer consisting of $MoSi_x$/Al. After standard chemical cleaning and an ozone treatment[5] to remove any polymerization products from the RIE step, 500 Å of TiW film containing approximately 20 at.% of Ti, according to RBS measurements, was sputtered deposited.

The commercial available CVD system, TEMPRESS, consists of a three zone horizontal furnace with a quartz tube 125 mm wide, 2100 mm long. The wafers are oriented parallel to the gas flow. All gases are admitted through massflow controllers. The system is pumped by a Roots blower package equipped with closed loop controlled nitrogen injection for pressure control.

Tungsten of the desired thickness was deposited by the reaction: $WF_6 + H_2$ ---> W + 6HF with argon gas as dilutant. The actual parameters were:

WF_6: 10 sccm H_2: 300 sccm Ar: 490 sccm
Temperature: 350 °C Pressure: 350 mtorr
Deposition rate: 120 Å/min.

After deposition, the tungsten and TiW layers were backetched with CF_4/O_2 in a RIE reactor operated at 13.56 Mhz using a laser endpoint detector.

To establish the performance of our CVD system, the dependence of the tungsten deposition rate on temperature for the WF_6-H_2-Ar system was examined using single-crystal p-type wafers (16-24 Ωcm) as substrates.

RESULTS AND DISCUSSIONS

It is known that the reaction between WF_6 and silicon substrate is a very fast process in kinetics and, thermodynamically, there is a significant free energy change for the reaction of $2WF_6 + 3Si \rightarrow 2W + 3SiF_4$[6]. In addition, the thickness of the self-limited tungsten layer depends on the deposition temperature[7].

In order to eliminate the influence of the reduction of WF_6 by the silicon substrates on the deposition rate, a continuous two step deposition process was used.

In the first step only Ar and WF_6 were admitted using the reduction between Si and WF_6 to form a self-limited layer of tungsten on the substrate. The parameters of this first step, i.e., temperature, WF_6 partial pressure, total gas flow rate and substrate pretreatment were kept constant in all the subsequent deposition rate studies to supply a tungsten layer of constant thickness prior to the deposition of tungsten via hydrogen reduction in the second step with argon as dilutant gas. Temperature was the only parameter varied from run to run. WF_6 and H_2 partial pressures, total gas flow and pressure were kept constant.

The curve fitting method was employed to deal with the total thickness of the tungsten from four-point probe measurements. During the curve fitting process, a tungsten layer of constant thickness Δ, obtained in the first deposition step, was substracted from the total thickness. If the reported activation energy of 0.71 eV/atom[8] is used, the resulting thickness for Δ is 126 Å, which is in good agreement with the reported value of 130 Å[8] for self-limiting tungsten films formed by silicon reduction of WF_6 and RBS measurements. Our results confirm that tungsten deposition in a WF_6-H_2-Ar system is controlled by a surface reaction. The deposition rate as a function of temperature can be expressed by Arrhenius' Equation, $r = A \cdot e^{-E/RT}$, on which the curve fitting procedure was based. On evaluation we obtained E = 0.71 eV/atom and $A = 3.6 \times 10^7$ Å/min (Fig. 1).

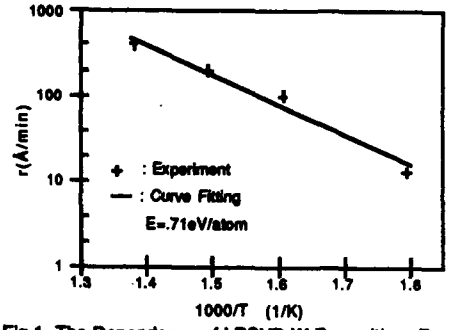


Fig.1 The Dependence of LPCVD W Deposition Rate on Temperature

We also tried to fill vias on wafers which did not receive the extra coating of sputtered TiW by selective tungsten deposition, using the above given parameters. For deposited films thicker than 4000 - 5000 Å, we noticed that selectivity was lost. Further, selective tungsten growth is inhibited by even very thin silicondioxide layers[9], resulting in a very unreliable process considering the large number and the varying depth of the vias to be filled. The deposition of a thin layer of TiW after sputteretching the vias will break up any thin silicondioxide layer that might be left, assuring good contact of the CVD tungsten to the underlying metal layer. No voids were observed when using this technique to fill the vias. Fig. 2 shows a 1.5 μm filled via after tungsten deposition.

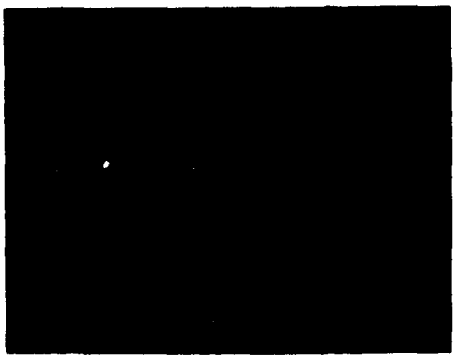


Fig.2 1.2 μm W deposited in a 0.8 μm deep via hole of 1.5 μm in diameter.

To reduce surface roughness effects during the backetch step, positive photoresist (AZ 1450 J) was spun onto the wafers and hardbaked (210 °C, 30 min.) before etching. Reactive ion etching in CF_4/O_2 plasma was utilized to backetch the resist/tungsten/TiW layer on the oxide surface. In Fig.3, the etchrate dependence of W, photoresist and SiO_2 on the percentage of oxygen in the gas mixture is shown. The endpoint was determined by the optical reflection method using a He-Ne laser. The output signal from the laser endpoint detector is shown in Fig. 4. Due to the slow etchrate of the underlaying material (SiO_2), the endpoint is very sharp. This method requires a rather well planarized oxide surface to avoid strings of W/TiW being left which could cause shorts. Figs. 5 and 6 illustrate backetched via holes ready for the second level metallization.

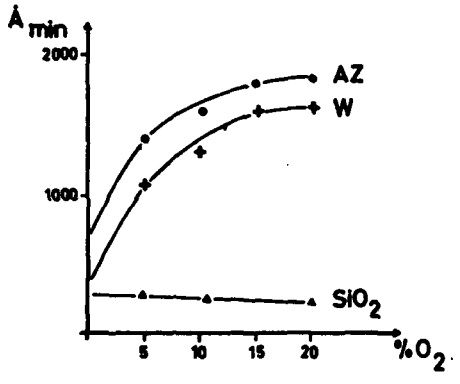


Fig.3 Reactive ion etch rates of W, AZ photoresist and SiO_2 as a function of the percentage of O_2 in CF_4/O_2. Pressure: 120 mtorr, R.F. power(13.56 MHz): 0.6 W/cm^2

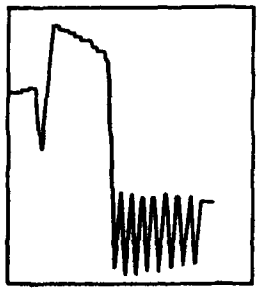


Fig.4 Backetch endpoint detection.



Fig.5 Backetched tungsten plug.



Fig.6 As Fig.5, seen from above.

Electrical measurements using six-terminal Kelvin structures[10] showed uniform specific contact resistances of $1.3\text{x}10^{-8} - 2.3\text{x}10^{-8}\ \Omega\cdot\text{cm}^2$.

The tungsten filled vias were tested in a commercially available multilayer CMOS process. Wafers with TiW/tungsten filled vias were compared with wafers using no via-filling and standard Al(Si) as second metallization layer. Electrical measurements gave low and reproducible contact resistance . Possible problems introduced by the backetch procedure, such as bridging in the second metallization layer due to unetched TiW/tungsten or interlayer isolation damage, did not occur.

CONCLUSION

We have shown that "selective" blanket deposition of CVD tungsten on sputtered TiW films with maintained selectivity against furnace tube and boat can be used to reliably fill vias for multilayer metallization systems.

REFERENCES

[1]. T. Moriya, S. Shima, Y. Hazuki, M. Chiba and M. Kashiwagi, *A Planar Metallization Process - Its Application to tri-level-Al-Interconnection*, IEDM Techn. Dig., pp. 550-553, 1983.

[2]. D. M. Brown, B. Gorowitz, P.A. Piacente, R.J. Saia, R.H. Wilson, D.W. Woodruff, *Selective CVD Tungsten via Plugs for Multilevel Metallization*, IEDM Techn. Dig., pp. 66-69, 1986.

[3]. Y. Pauleau, *Chemical Vapor Deposition of Tungsten Films for Metallization of Integrated Curcuits*, Thin Solid Films, 122, 243(1984).

[4]. R.A. Levy, M.L. Green, P.K. Gallagher and Y.S. Ali, *Selective LPCVD Tungsten for Contact Barrier Applications*, J. Electrochem. Soc., 133, 1905(1986).

[5]. H. Norström, M. Östling, R. Buchta and C.S. Petersson, *Dry Cleaning of Contact Holes Using Ultraviolet (UV) Generated Ozone*, J. Electrochem. Soc., 132, 2285(1985).

[6]. J.-S. Lo, R.W. Haskell, J.Gerald Byrne and A. Sosin, *A CVD Study of the Tungsten-Silicon System*, Proceedings of the Fourth International Conference on CVD, 1973, pp. 74-90.

[7]. K. Y. Tsao and H. H. Busta, *Low Pressure CVD of Tungsten on Polycrystalline and Single Crystal Silicon via Silicon Reduction*, J. Electrochem. Soc., 131, 2702(1984).

[8]. E. K. Broadbent, C. L. Ramiller, *Selective Low Pressure Chemical Vapor Deposition of Tungsten*, J. Electrochem. Soc., 131, 1427(1984).

[9]. H. H. Busta and C. H. Tang, *Film Thickness Dependence of Silicon Reduced LPCVD Tungsten on Native Oxide Thickness*, J. Electrochem. Soc., 133, 1195(1986).

[10]. S. J. Proctor, L. W. Linholm, J.A. Mazer, *Direct Measurements of Interfacial Contact Resistance, End Contact Resistance, and Interfacial Contact Layer Uniformity*, Trans. El. Dev., ED-30, 1535(1983).

Al/TiN/$TiSi_2$ CONTACTS TO ULTRA SHALLOW JUNCTIONS

E. Ling, H.S.Gamble, B.M. Armstrong and J.H. Montgomery

Department of Electrical and Electronic Engineering,
Queen's University of Belfast, Ashby Building,
Belfast BT9 5AH N. Ireland

Processes for the production of reliable aluminium contacts to 110nM junctions are described. In the case of P^+-N junctions the temperature-time cycles have been reduced to minimise redistribution of the boron impurities. A plasma enhanced rapid thermal processor has been employed.

1. INTRODUCTION

Ultra Large Scale Integration requires very shallow junctions of the order of 100nM. These are readily achieved with ion implantation for arsenic N^+ junctions, but boron P^+ junctions are more difficult. The low energies required for boron implantation leads to considerable ion-channelling and the profile tail can contribute significantly to the junction depth. Implanting through a metal layer or using molecular BF_2 implants with or without preamorphizing the silicon are possible alternatives. However, it is difficult to achieve an activation anneal sufficient to provide good reverse bias characteristics without increasing the junction depth. As the junction depth is reduced this problem is exacerbated and at depths of 100nM the reverse bias leakage is as high as $10^{-7}A.cm^{-2}$ [1].

Alternatives to ion implantation for shallow junction formation are presently being persued [2]. A laser doping technique has been used to obtain a value of $47nA.cm^{-2}$ at 10v for a 100nM junction with a peak boron concentration of $2 \times 10^{20}cm^{-3}$ and a sheet resistance of 150ohms/square. The rapid thermal diffusion of boron from an oxide source has also shown considerable improvement in reverse bias characteristics of shallow junctions, without compromising surface concentration or sheet resistance [3].

Having obtained the ultra shallow junctions it is necessary to provide them with stable, reliable contacts. This is a major problem since aluminium reacts with and interdiffuses into the silicon substrate in the contact window. Therefore if aluminium is to be used for the metal interconnect a second metal that acts as a diffusion barrier must be sandwiched between the aluminium and the silicon.

Self-aligning titanium disilicide to the shallow junctions helps to reduce interconnect and contact resistances. Titanium disilicide however does not act as a diffusion barrier to silicon and aluminium and hence there is still a need for a barrier metal. Of the barrier metals proposed in the literature titanium nitride appears to be the most compatible for use with a titanium disilicide layer since the titanium disilicide surface can be converted to TiN by a high temperature anneal in a nitrogen ambient.

This paper will describe various technologies for providing reliable stable contacts to silicided shallow junctions. The shallow junctions were produced by the furnace diffusion of arsenic from a spin on source for the N^+ junctions and by the rapid thermal diffusion of boron again from a spin on source for the P^+ junctions.

2. DEVICE STRUCTURE

Three photographic masks were used in the

manufacture of the diode structures. The designs included a circular geometry device, and five rectangular devices with different perimeter lengths. All these structures had approximately the same junction area of 1.8 x $10^{-4}cm^2$. The aluminium over the contact windows was smaller than the diffused area to avoid any field plate effects during reverse bias leakage measurements. The incorporation of these test structures on the mask set permits measurement of the contributions to current flow due to area, perimeter and corners. An array structure with 30 times the area and a perimeter length of 4.8cms is also included to check the parameter values deduced.

3. SAMPLE PREPARATION

3.1. Shallow boron junctions

These devices were fabricated on n type silicon wafers of <100> orientation and a dopant density of 0.8 - 1.0 x $10^{15}cm^{-3}$ phosphorus atoms. Prior to oxidation these wafers were mechanically back damaged to aid in relieving the point defects and structure imperfections by acting as a sink for self-interstitials and metallic impurities. This ensures a high carrier lifetime in the bulk silicon and hence minimises reverse bias leakage currents.

The field oxide was grown in two stages. The initial oxide was grown at a temperature of 1050C for 90 minutes in a dry oxygen and trichloroethylene ambient. This step was used to minimise the sodium content in the oxide and the redistribution of phosphorus in the silicon substrate. The oxide was removed from the back of the wafer and a phosphorus source diffusion carried out for 15 minutes at 1000C from a solid phosphorus source. This back diffused layer also helps to getter out impurities and provides for a good back contact. The field oxide thickness is increased by a wet oxidation step at 1000C to 0.3μm.

Emulsitone Borofilm 100 solution was used as the spin on source. After curing the boron was driven in by heating the silicon wafer in a nitrogen ambient to a temperature of 1000C for 3 minutes with incoherent light. Subsequent processing was standard. A typical profile for the boron junction is shown in Figure 1.

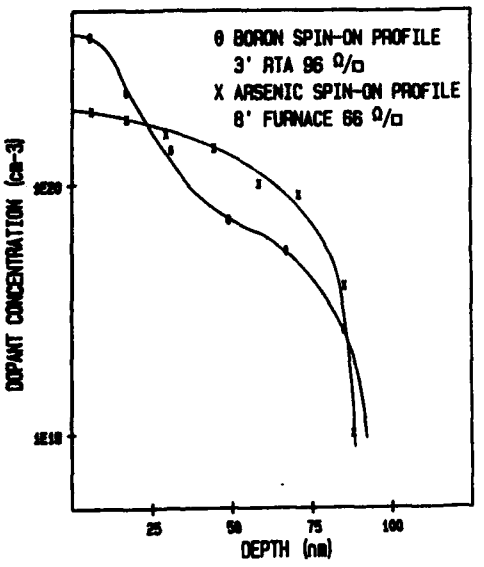


FIGURE 1

The dopant distribution for both the Arsenic and Boron shallow junctions.

3.2 Shallow arsenic junctions

These devices were fabricated on p type silicon wafers of <100> orientation and a dopant density of $10^{16}cm^{-3}$ boron atoms. The wafers were subjected to the same back damaging and gettering schedule described for the boron junctions. The process was modified to permit field threshold voltage adjustment. The LOCOS process was therefore employed and a spin on dopant layer, with boron concentration of $10^{19}cm^{-3}$, was used to enhance the substrate doping in the field regions. The oxidation process, again undertaken in two stages yielded an oxide thickness of 0.7μm and a field threshold voltage of 20v.

Emulsitone Arsenic Film 3739N was used as the source of Arsenic for the N^+-P junction.

The arsenic was driven in for 8 minutes at 1000C by standard furnace diffusion and the processing was standard. A typical profile for the Arsenic junction is also included in Figure 1.

4. LOW LEAKAGE JUNCTIONS

Typical leakage current density in the range 1-4 x 10^{-9}A.cm^{-2} at a 3 volt reverse bias was measured for the test structures. Separating the current components into area, perimeter and corners the best fit perameters were 103pA.cm^{-2} for junction area, 2.5pA.cm^{-1} for junction perimeter and a negligible contribution due to each corner. A discontinuity in current characteristics was observed in some diodes at bias voltages around 3.5v, possibly due to local generation sites.

At 3v reverse bias the depletion layer width in the substrate is 2.1μm. Assuming the bulk generation current is dominated by the depletion region then the lifetime in the bulk silicon is 2.4mSec. From the perimeter contribution to current flow the surface recombination velocity S_o is 12cm.sec^{-1}. Low leakage properties have also been observed for the shallow arsenic junctions. Typical current densities at a reverse bias of 3v were of the order 5 x 10^{-8}A.cm^{-2}.

5. TITANIUM DISILICIDE

The titanium disilicide layers employed in this work were produced by the sputter deposition of titanium and a two stage silicide formation process in a low pressure rapid thermal processing chamber. For the 100nM junctions employed a practical limitation on silicide thickness exists. The influence of silicide thickness on reverse bias leakage current density for the shallow boron junctions is shown in Figure 2. For thicknesses up to approximately 40nM there is little dependance of leakage current on silicide thickness. However for the 60nM thickness there is a very large increase in leakage current. This is due to the fact that the junction is now operating as a Schottky barrier diode. The onset of Schottky action has been established to be between thicknesses of 55nM and 60nM of titanium discilicide on the 100nM boron junctions [4].

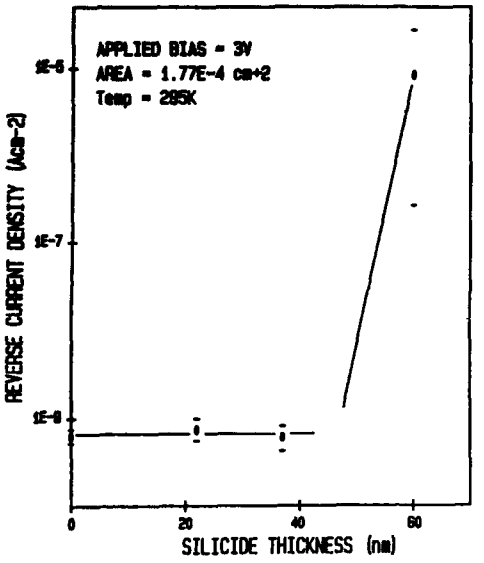


FIGURE 2

Reverse current density versus titanium disilicide thickness at a reverse bias of 3v.

6. TiN FORMATION

The initial results were achieved by the conversion of the $TiSi_2$ surface layers to TiN by thermal annealing in an ammonia ambient. This was accomplished in a rapid thermal processor at a temperature of 1000C for 15 seconds. All other aspects of the process were as previously described. The reverse bias leakage currents of the structures were measured. It is important to note that these characteristics did not exhibit the discontinuities at reverse bias of 3.5 volts previously described for the junctions formed without TiN. This can be attributed to the approximate 10% increase in junction depth caused by the additional processing to form

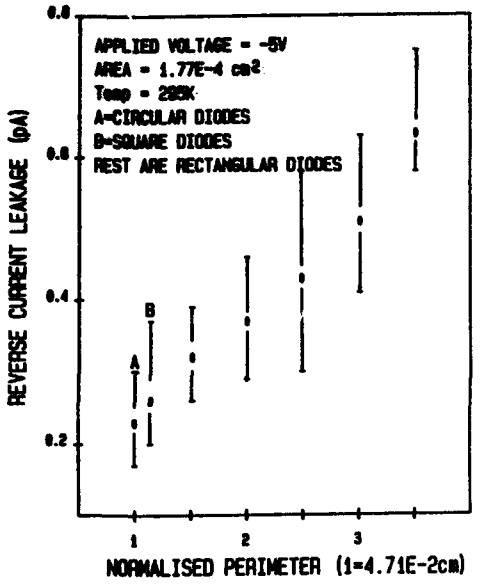


FIGURE 3

Reverse leakage current versus perimeter length at an applied voltage of 5v. Contact system Al-TiN-$TiSi_2$.

TiN. The leakage current measured at 5v versus junction perimeter is given in Figure 3. Leakage current is again dominated by carrier generation at the oxide-silicon interface and so is almost linear with perimeter length. The surface recombination velocity S_o is 15cm.sec^{-1} and the lifetime in the bulk silicon is 1mSec. The technology employed has not therefore impaired the high quality of these shallow junctions.

It is desireable to minimise the amount of TiN required for the barrier layers since its resistivity is much higher than for $TiSi_2$. A temperature-time cycle sufficient to form the desired TiN thickness with minimum redistribution of the dopant impurities is also required. Three alternatives to the formation of TiN have been investigated.

The conversion of titanium to TiN has been performed in an ammonia plasma at temperatures around 350C. While this offers minimised redistribution of impurities and good control of ultimate TiN thickness, it does not yield self-alignment to the underlying contact areas and is therefore less desirable for ULSI.

The conversion of fully formed $TiSi_2$ in an ammonia atmosphere with and without plasma excitation has been performed in a rapid thermal processor. This produces TiN self-aligned the silicided contacts and ensures minimised drive in of the junctions. There are two possible options in this approach. The TiN thickness to ensure junction integrity after aluminisation and annealing can be formed. Alternatively a thin TiN layer can be formed. This acts as an oxidation barrier on the silicide contact and the TiN layer can then be increased to the desired thickness by reactive sputtering followed by aluminisation. The deposited TiN layer is then patterned using the aluminium mask.

Titanium silicide is normally formed by a two stage rapid thermal process. The first low temperature step ensures self-alignment to exposed silicon while the second converts this silicide to its high conductivity form. The second annealing step was adapted to incorporate an ammonia atmosphere in the processor thus achieving simultaneously the high conductivity $TiSi_2$ together with the required TiN surface layers.

The results obtained using these techniques will be presented. The layers were subjected to post aluminium anneals at 450°C and 500°C and the criteria for aluminium penetrations taken as an order to magnitude increase in leakage current at 3v.

REFERENCES

[1] Wilson. R.G. J. Appl. Phys. (1983) 6879.

[2] Carey, P.G., Sigmon, T.N., Preso, R.L. and Fahlen, T.S. IEEE Trans. Electron. Dev. Lett. (1985) 291.

[3] Ling, E., Maguire, P., Gamble, H.S. and Armstrong, B.M. Proc. ESSDERC (1986).

[4] Ruddell, F., Ling, E., Armstrong, B.M., Gamble, H.S. and Raza, S.H. Shallow Junctions Contacts for Latch up Resistance in CMOS, this volume.

THE EFFECT OF ION-IRRADIATION AND RAPID THERMAL ANNEALING ON $TiSi_2$ AND $MoSi_2$

Leif GRANBERG*, Christos KRONTIRAS**, Jaakko SAARILAHTI* and Ilkka SUNI*

* Technical Research Centre of Finland, Semiconductor Laboratory, Otakaari 7 B, SF-02150 Espoo, Finland

** Department of Physics, University of Patras, Patra 26-110, Greece

The effects of ion-irradiation and subsequent rapid thermal annealing have been studied in thin $TiSi_2$ and $MoSi_2$ films. A substantial increase in the film resistivities is observed for As^+ implantations up to a dose of 10^{14} ions/cm^2. An almost complete recovery of the ion induced damage is obtained in $TiSi_2$ after rapid thermal annealing for 10 sec in the 500 to 600°C range. In the case of $MoSi_2$ 10 sec at 1100°C is required to attain the initial resistivity level.

1 INTRODUCTION

The shrinkage in silicon integrated circuit dimensions has led to considering thin films of refractory disilicides for low resistivity gates and interconnections. Co-deposited silicides are typically highly disordered or amorphous in their as-deposited state. Consequently, they exhibit unacceptably high resistivities and require annealing at elevated temperatures. During subsequent processing steps the formation of self-aligned MOSFETs might expose the silicide film to ion-irradiation, which induces damage in the film and thereby an increase in the film resistance. Previous studies have established metal rich silicides as radiation hard materials whereas many disilicides are relatively radiation soft [1]. Heat treatments, such as required by source and drain implant activation, can be expected to remove at least part of the damage. However, many high melting point silicides that become amorphous at a relatively low ion dose need high annealing temperatures for complete recovery. Pursuant to an earlier report on the transport properties of ion-irradiated silicides [2] we report here on the effects of arsenic implantation and rapid thermal annealing in thin films of $TiSi_2$ and $MoSi_2$.

2 EXPERIMENTAL

The silicide layers were formed by solid phase reactions of sputtered metal films with silicon. $TiSi_2$ was formed by sputtering 370 A of Ti followed by 120 A of amorphous silicon on n-type <100> Si wafers having a room temperature resistivity of 2...4 Ωcm. The wafers were then annealed in vacuum using a two step annealing procedure, first for 20 min at about 450°C and then for 30 min at about 580°C. $MoSi_2$ was formed by sputtering 270 A of Mo on undoped polycrystalline silicon deposited on oxidized silicon wafers. Annealing was carried out in vacuum at 680°C for up to 54 min. The resulting layer thicknesses were 930 A ($TiSi_2$) and 700 A ($MoSi_2$) as determined from RBS measurements. Some of the wafers were then implanted at room temperature with arsenic to doses in the range $N_s = 10^{13}...10^{15}$ cm^{-2}. The ion energies, 150 keV for $TiSi_2$ and 190 keV for $MoSi_2$, were selected to match the damage depth with the film thickness. The TRIM 86 simulation program was used for range calculations [3]. To avoid beam heating the ion current was held below 20 μA. The recovery of the damage was investigated by subjecting the samples to rapid thermal annealing for 10 to 30 s in the temperature range of 300...1150°C using a Heatpulse quartz halogen

lamp system. The resistivities of the silicides were measured in the van der Pauw configuration [4] using a cryostat to control the sample temperature between 4.2 K and 300 K.

3 RESULTS AND DISCUSSION

The effects of ion implantation on the residual resistivity of $TiSi_x$ are shown in fig. 1. The resistivities of an unimplanted sample were 1.5 μΩcm at 4.2 K and 13.1 μΩcm at 300 K. After an arsenic dose of $N_{As} = 10^{16}$ cm^{-2} the resistivities at 4.2 K and 300 K are almost identical in the order of 150 μΩcm. In $TiSi_x$ films irradiated at high dose levels, at room temperature part of the current is carried through the substrate as deduced from the temperature dependence of resistivity. The residual resistivity ρ_0 (at 4.2 K) is almost linearly proportional to log N_{As}. The temperature dependence of resistivity (the temperature coefficient is positive even for the highest irradiation dose) seems to indicate that $TiSi_x$ is not fully amorphized even at an As^+ dose of 1×10^{16} cm^{-2}.

The corresponding behavior of $MoSi_x$ is traced in fig. 2. In this case the unimplanted samples have a resistivity of 77 μΩcm at 4.2 K and 91 μΩcm at 300 K. The resistivity increases rapidly with increasing arsenic dose and peaks at $N_{As} \approx 10^{15}$ cm^{-2}. The temperature dependence of resistivity becomes strongly negative at high implantation doses. The resistivities corresponding to the implantation dose of $N_{As} = 10^{15}$ cm^{-2} are 1308 μΩcm at 4.2 K and 879 μΩcm at 300 K. The maximum that appears in the dose dependence of the resistivity is analogous to the resistivity maximum observed for thermal annealing of co-deposited $MoSi_x$. There, as in the case of WSi_x, the behavior is attributed to the transition from the hexagonal low-temperature phase to the high-temperature tetragonal phase [5]. Nucleation of the hexagonal phase of $MoSi_x$ under ion-irradiation has been previously reported [6]. It is not clear, however, that also the stacking faults responsible for the increased resistivity in the case of thermal annealing should nucleate from thermal spikes induced by energetic ions.

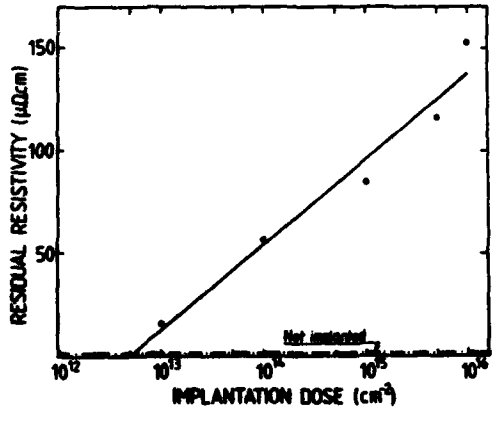


Fig. 1 Residual resistivity vs. dose for As^+ irradiated $TiSi_x$ films. The irradiation energy energy was 150 keV. Also indicated is the resistivity level of a non-irradiated film.

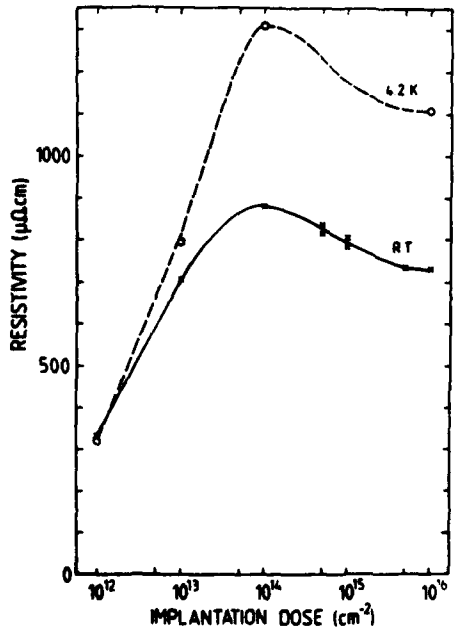


Fig. 2 Room temperature and residual resistivities vs. dose for As^+ irradiated $MoSi_x$ films. The irradiation energy was 190 keV.

The effects of isochronal rapid thermal annealing on the implanted $TiSi_2$ are depicted in fig. 3 for two different ion doses. The samples implanted at $N_d = 10^{15}$ cm^{-2} recover almost completely after 10 s at 500 °C. For $N_d = 10^{16}$ cm^{-2}, the minimum temperature for the recovery is 600 °C, but annealing even at 900 °C still leaves the room temperature resistivity 40 % higher than in the unimplanted case. Also the residual resistivity is twice the initial value suggesting that the ion damage is only partially removed. The dependence of the resistivity on the annealing time, shown in fig. 4, is not very strong below 950°C. At higher temperatures, however, the resistivity starts increasing and cracks are observed by visual inspection. Other investigators have reported that furnace annealing of thin $TiSi_2$ films above 900°C can lead to serious morphological and electrical instabilities above 900°C [7]. The same instabilities are observed here for an extended duration of halogen lamp annealing.

The cases of $MoSi_2$ irradiated with two different doses are depicted in fig. 5. It is worth noting that the recovery is faster in the $N_d = 10^{16}$ cm^{-2} case than in the sample corresponding to the resistivity maximum at $N_d = 10^{15}$ cm^{-2}. For periods longer than 10 s the resistivity is not very sensitive to the annealing time. The difference in the annealing behavior of the two samples with different implantation doses sug-

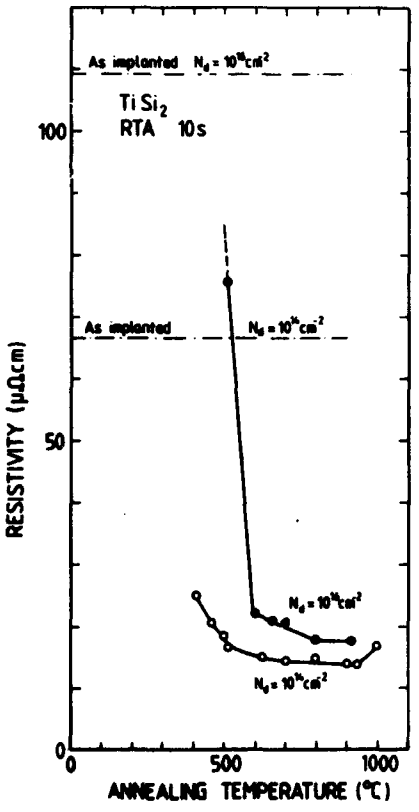


Fig. 3 Resistivity vs. annealing temperature for $TiSi_2$ films irradiated with two different doses. Also indicated are the as-implanted resistivity values.

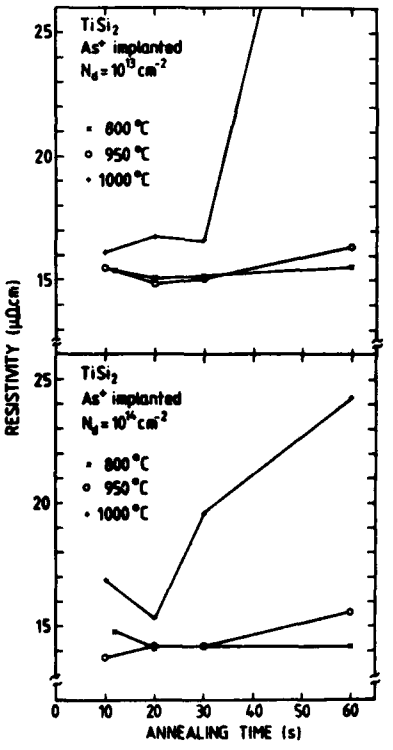


Fig. 4 Resistivity vs. annealing time for $TiSi_2$ films irradiated with two different doses. The resistivity increase for RTA at 1000°C coincides with cracks observed by optical microscopy.

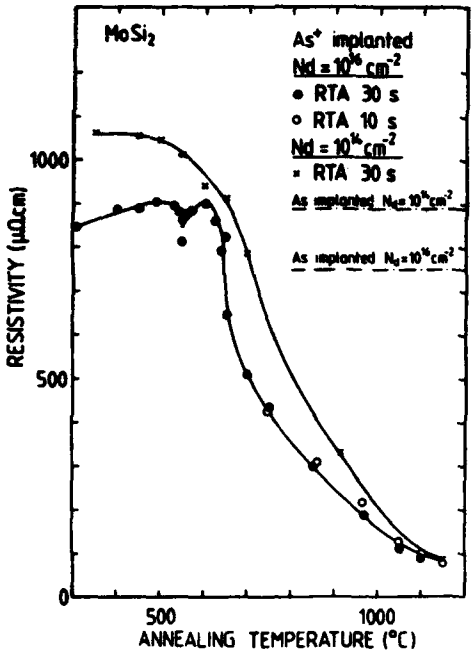


Fig. 5 Resistivity vs. annealing temperature for $MoSi_{2}$ films. As-implanted values are indicated by the dashed/dotted lines.

gests that after low-dose irradiation the recrystallization proceeds directly to the tetragonal $MoSi_{2}$ phase whereas the high-dose implanted samples undergo a transition from the hexagonal to the tetragonal stacking order. It appears that for $MoSi_{2}$ a short annealing step above 1000°C would yield the optimum conditions for damage recovery.

4 CONCLUSIONS

We have investigated the annealing behavior of thin $TiSi_{2}$ and $MoSi_{2}$ films after high dose arsenic implantations. The residual resistivities in these silicides increased by 1 to 2 orders of magnitude for the highest implantation dose of 1×10^{16} cm^{-2} indicating a high degree of disorder. Most of the damage was removed by rapid thermal annealing for 10 sec at 600°C ($TiSi_{2}$) or 1100°C ($MoSi_{2}$) as determined by electrical resistivity measurements.

REFERENCES

[1] Hewett C.A, Suni I., Lau S.S., Hung L.S., and Scott D.M., in Ion Implantation and Ion Beam Processing of Materials, G.K. Hubler, O.W. Holland, C.R. Clayton, C.W. White, Eds., (North-Holland, N.Y. 1984), MRS Symposia Proceedings, Vol.27, p.145

[2] Suni I., Grönberg L., and Saarilahti J., Le vide - les couches minces, 42/236 (1987) 233

[3] Ziegler J.F., Biersack J.P., Littmark U., The Stopping and Ranges of Ions in Matter (Pergamon Press, New York, 1985)

[4] Van der Pauw L.J., Philips Research Reports, 13 (1958) 1

[5] D'Heurle F.M., LeGoues F.K., Joshi R., and Suni I., Appl. Phys. Lett. 48 (1986) 332

[6] D'Heurle F.M., Petersson C.S., and Tsai M.Y., J. Appl. Phys. 51 (1980) 5976

[7] Ting C.Y., d'Heurle F.M., Iyer S.S., and Fryer P.M., Electrochemical Society, Extended Abstracts 85-1 (1985) 387

EVALUATION OF ELECTROMIGRATION ACTIVATION ENERGY BY MEANS OF NOISE MEASUREMENTS AND MTF TESTS

A. Diligenti, P.E. Bagnoli, B. Neri.

Istituto di Elettronica e Telecomunicazioni, Facoltà di Ingegneria, Università di Pisa, Via Diotisalvi, 2, I-56100 Pisa, Italy

G. Specchiulli

TELETTRA S.p.A., Quality and Reliability Dept., Via Trento 30, I-20059 Vimercate (Milano), Italy

ABSTRACT

Activation energy E_a of grain-boundary vacancies was evaluated by means of noise measurements and MTF tests in narrow Al/Si (1%) resistors. The values obtained by these techniques are 0.93 and 0.96 eV respectively.
Noise measurements revealed that after every temperature change the microstructure of the films was unstable. The presence of instabilities can strongly affect the E_a value.

1. INTRODUCTION

Low-frequency (10^{-2} - 2 Hz) noise measurements have been proposed as a new, sensitive technique able not only to detect electromigration (EM) under non-accelerated test conditions, but also to give a value of E_a, activation energy of grain-boundary vacancies, characteristic of the first stage of the EM phenomenon |1|, |2|. Low-frequency spectra, whose dependence on frequency f follows a law $1/f^{\gamma}$ with $2 < \gamma < 2.8$, are caused by vacancy generation-recombination processes and their value at a given frequency f_o, when $\gamma = 2$, is assumed to be proportional to the vacancy flux J_v according to the relationship

$$S_R(f_o) = \frac{B}{f_o^2} J_v = B'N(eZ^*/KT)\rho j D_o e^{-E_a/KT} \qquad (1)$$

where ρ is the metal resistivity, j the current density, N the atom density, eZ* an effective ion charge and D_o the infinite temperature diffusion coefficient; K and T have the usual meaning and B' is a suitable proporzionality constant. S_R is the power spectral density of resistance fluctuations, that is $S_R(f) = S_V(f)/I_o^2$, where $S_V(f)$ is the power spectrum of voltage fluctuations across the resistors supplied with a constant current I_o.

Owing to the relationship (1), it is a straight forward procedure to deduce E_a from measurements of S_R at various current densities and temperatures |2|.

In this work we attempted, for the first time, to compare the results obtained by applying noise and MTF techniques to VLSI resistors. The nominal width of the Al/Si(1%) resistors employed is 2 and 4 μm, with a cross section of 1.6 and 4 μm^2 respectively. However the actual width of 2 and 4 μm resistors are about 1.4 and 3.5 μm respectively. The sample fabrication procedure is described in another paper presented at this conference |3|.

Often a bamboo structure occurs in the samples |3| and this fact is probably the main cause of the strong dispersion of results (times to failure and activation energies) obtained from noise measurements and MTF tests.

2. EXPERIMENTS

Before noise measurements, the resistance vs temperature behaviour of all samples was determined in the range 50-150°C by putting them in a thermostatic silicone-oil bath and

then recording current, voltage (four contacts test-pattern) and temperature simultaneously. The sample to be tested was put in a bridge configuration (all the resistors used save the sample are flicker noise free) and then supplied with the proper current density by means of Pb batteries. The temperature of the resistor depends on the current, on the thermal conductivity of the package and on the environmental temperature. The sample, the current supply system and the purpose-designed amplifier |2| were shielded electrically and magnetically. The amplified signal was processed by an HP 3561A dynamic signal analyzer.

A first set (S # 1) of 16 resistors (2 μm wide) were all tested at the same nominal current density (4.8×10^{6} A/cm^2), the temperature of the samples lying in the range 63-73°C because of small differences in the thermal conductivity of packages and slightly different room temperatures. In spite of nearly equivalent test conditions, noise spectra displayed a random behaviour, both in magnitude and in γ values. We found γ values in the range 2-2.8 and S_R(20 mHz) in the range 1.69×10^{-13}-8.45×10^{-11} Ω^2/Hz, whereas for wider resistors (14 μm) such a great dispersion was never observed |4|. The samples were then subjected to an MTF test under the following conditions: $J=3.2\times10^{6}$ A/cm^2 and T=205°C |3|. During the test, the resistance variation vs time was also recorded. The aim of performing noise measurements and MTF test on the same samples was that of finding a correlation, if any, between EM noise and time to failure.

A second set of resistors 2μm and 4μm wide (S#2) was characterized only by means of noise measurements in order to find the value of E_a.

During this kind of test, which requires many S_R measurements on the same sample at various j and T, the following facts were observed: i) the spectrum value, at a given frequency and for fixed j and T, is generally a decreasing function of the time, as if the film microstructure were undergoing a sort of "current annealing" which tends to reduce the causes of EM noise; only after many hours (about 4) the spectrum becomes stationary; ii) when EM occur, the signal from the amplifier consists of a succession of pulses as shown in fig. 1. Similar pulses have already been observed in Al samples during thermal cycling |5|.

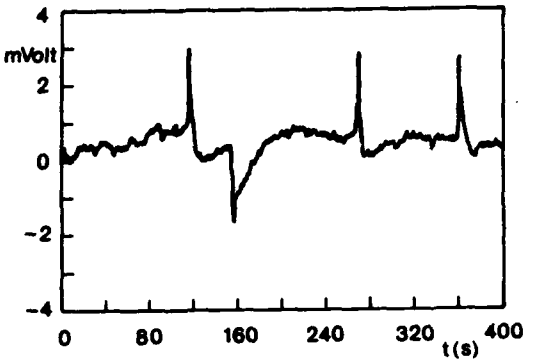


FIGURE 1

Random pulses at the output of the amplifier (gain:10^4) in 2 μm resistor above threshold current density.

The amplitude and frequency of the pulses are increasing functions of j and T but, for fixed j and T, both amplitude and frequency decrease with time; iii) a threshold value of the current density j_T ($j_T \simeq 3.5\times10^{6}$ A/cm^2), which is the same for 2 and 4 μm resistors, seems to exist for the appearance of the pulses. The minimum pulse amplitude is equivalent to a percentage resistance variation of about 0.1 ppm.

The evaluation of E_a was made only for 4 μm resistors because the others (2 μm) showed too high an instability of the spectra during the observation time. E_a is deduced from Arrhenius plots obtained by putting the relationship (1) in the form

$$\ln S_F(f_o) = \ln \frac{S_R(f_o)T}{\rho j} = \ln \bar{B} - \frac{E_a}{KT} \qquad (2)$$

where $\bar{B} = \frac{B'NeZ^*}{K} D_o$

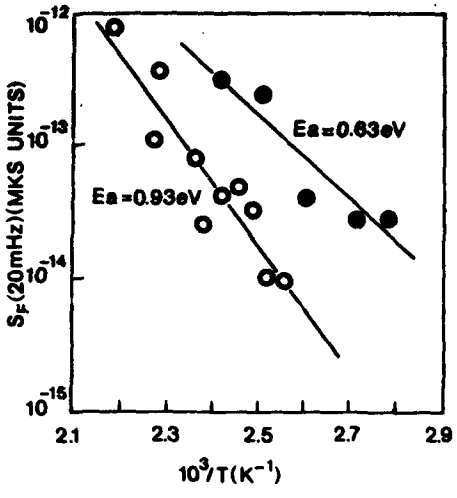


FIGURE 2

Arrhenius plots for 4 μm resistors: from "fast" noise measurements (black dots) and from stationary spectra (open dots).

Fig. 2 shows two examples of these plots. The first one (black dots) refers to a sample on which a set of *fast* noise measurements was performed. The word fast means that the noise spectra were taken after the sample reached thermal equilibrium, but without waiting for the spectra to attain a stationary behaviour (see point i). In this case the slope of the straight line gives E_a=0.63 eV.

The second plot (open dots) is obtained from another sample, for which each S_R value was taken when the spectrum became stationary. In this case E_a=0.93 eV. It must be noted that this sample, when previously subjected to a set of fast measurements, also displayed strong instabilities in spectrum values.

E_a was also deduced from MTF test performed at three different temperatures, 205, 223, 261 °C on S # 1 and the value obtained is 0.96 eV |3|.

3. DISCUSSION

A comparison between the results of noise measurements and MTF test performed on S # 1, did not reveal any certain correlation among the information extracted using the two techniques. It must be pointed out that even for the MTF test alone there is no correlation between the time to failure (TTF) and the resistance variation or the occurrence of the first spike |3|. For such narrow stripes in fact, the resistance variation vs time did not show a regular behaviour but the resistance increase (decrease was sometimes observed) is characterized by spikes randomly distributed in time |3|. Moreover the standard deviation of the TTF distribution is higher than for wider stripes (σ =1.3).This anomalous behaviour shown by samples subjected to a traditional test corresponds to their anomalous behaviour as far as the noise measurements are concerned. The instability of the spectra can, perhaps, reflect a non-equilibrium situation in the microstructure of the films. At this stage of the work we can only suppose a redistribution of silicon at grain-boundaries caused by the temperature and the current, and/or the annihilation of defects such as dislocations |5|.

As far as the activation energy E_a is concerned, the value of 0.96 eV found by means of MTF tests is in agreement with that obtained using the noise technique for wider samples |4|. In applying this technique to 4 μ m samples, the long time necessary to obtain the points of "non-fast"plots(fig.2),did not allow us to test a sufficient number of resistors to give a reliable value for E_a, but the preliminary value obtained, that is 0.93 eV, is in good agreement with the preceding ones.

4. ACKNOWLEDGEMENTS

This work was partly supported by Ministero della Pubblica Istruzione of Italy.The authors wish to thank C. Caprile for sample preparation.

REFERENCES

|1| Diligenti, A., Neri, B., Bagnoli, P.E., Barsanti, A. and Rizzo, M., IEEE Electron Device Letters, EDL-6 (1985) 606.

|2| Neri, B., Diligenti, A. and Bagnoli, P.E., IEEE Trans. Electron Devices, in print.

|3| Specchiulli, G., Fantini, F. and De Santi, G., Electromigration in narrow Al(Si) stripes for VLSI: comparison between MTF and resistometric methods for life predictions, this volume.

|4| Bagnoli, P.E., Ciucci, S., Diligenti, A. and Neri, B., in 3rd Int. Conf. on Quality in Electronic Components (Bordeaux, 1987) 62-68.

|5| Bertotti, G., Celasco, M., Fiorillo, F. and Mazzetti, P., J. Appl. Phys., 50 (1979) 6948.

ELECTROMIGRATION IN NARROW Al(Si) STRIPES FOR VLSI: COMPARISON BETWEEN MTF AND RESISTOMETRIC METHODS FOR LIFE PREDICTIONS§

G. SPECCHIULLI, F. FANTINI*

Telettra S.p.A., Quality and Reliability Dept. - 20059 Vimercate-Italy

G. DE SANTI

SGS Microelettronica S.p.A. - 20041 Agrate-Italy

Al (1% Si) electromigration in narrow stripes was studied by means of MFT and resistometric methods. For very narrow stripes no correlation could be found between the variation of resistance and the end of life. A simple model was developed to explain the observed small resistance variation.

1. INTRODUCTION

Aluminium electromigration (EM) is one of the more dangerous failure mechanisms in silicon integrated circuits and the connected risks are increased by the continuous scaling of dimensions, which causes higher current density within the stripes [1].

The expected life of the stripes in real operating conditions can be predicted by extrapolating the results of accelerated stress tests, by using the Median Time to Failure (MTF) method [2]. However this approach demands quite long test times, so that quicker methods, which try to infer the stripe life from resistance variarion, ΔR, at the beginning of the stress, have been developed [3-6]. However a clear correlation between the results of quicker ΔR tests and life tests has not yet been established experimentally [7], in particular for very narrow stripes suitable for VLSI application.

The purpose of this work is to verify whether such a correlation exists, by monitoring ΔR and final opening during accelerated life tests performed on 1.4 μm wide Al (1% Si) stripes.

2. TEST SET-UP AND SAMPLES

The experiments were performed by using hot plates with very high thermal capacity and accurate temperature control (±1°C up to 250°C). A fully automatized test set was designed, which enables a continuous resistance measurement and records the time to failure (TTF). All the data are stored in a mass memory and are automatically elaborated. A detailed description of the test set is reported in [8].

The samples are Al-Si stripes deposited by sputtering from Al/1% Si target at 150°C, over a double layer of 700 nm of thermal SiO_2 and of 900 nm of chemical vapor deposited (CVD) posphosilicate glass (PSG, with 12% P). The metal film was patterned by reactive ion etching, and annealed in pure Hydrogen at 450°C for 30 minutes, finally the metallization was covered with 700 nm of CVD PSG (5% P).

Different kinds of test patterns are contained in the chips, including stripes of various width and length, all provided with Kelvin contacts to measure the voltage drop under high current stress.

---

*Present address: Scuola Superiore di Studi Universitari e di Perfezionamento S. Anna, 56100 Pisa - Italy

§Paper partly supported by ECC under ESPRIT contract 554 and by CNR under Progetto Finalizzato MADESS.

3. EXPERIMENTS

This paper deals only with the results of the 1.4 μm wide, 160 μm long stripes, tested at actual temperatures of 261°C, 223°C and 205°C, with a current density of 3.2×10^6 A/cm^2. A total of 80 stripes were submitted to test: 20 at each of the extreme temperatures and 40 at the middle temperature.

The intrinsic differences among the stripes generally cause a split in the actual stress conditions: different test temperatures due to Joule heating and to resistance increase during the test, different current densities due to geometrical differences. The increase of temperature was taken into account by evaluating the temperature coefficient, α, and the total thermal resistance, R_{th}, of each stripe be measuring the resistance as a function of the current (5÷40 mA) at different temperatures (25÷250°C) [9]. The actual current densities were estimated by calculating the real cross section of each stripe. These characteristics of the test stripes are reported in Table I. This set of correction rules enable us to normalize the recorded TTF to reference stress conditions, by using the Black formula [2], with n=1.

4. RESULTS

Fig. 1 reports the log-normal plots, evaluated by means of the normalized TTF's, for the three tests performed; the estimations of the MTF's, standard deviations and correlation coefficients are reported in Table II.

We also recorded the 80 resistance diagraphs; figures 2, 3 and 4 are significant examples of such plots.

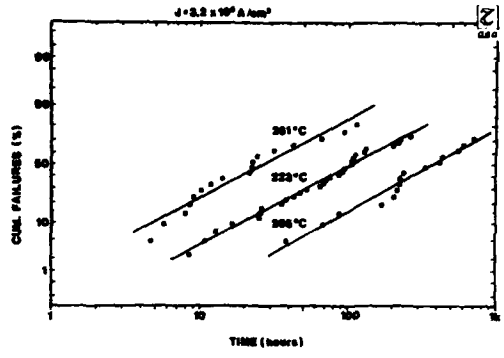


Fig. 1 Lognormal plots of the TTF obtained in the accelerated tests. Failure criterion was the open circuit.

TABLE II - Results of stress tests

Test temp.	MTF [h]	σ	Corr. coeff.
261°C	26.82	1.3448	0.98841
223°C	105.24	1.3448	0.99252
205°C	373.67	1.2823	0.98208

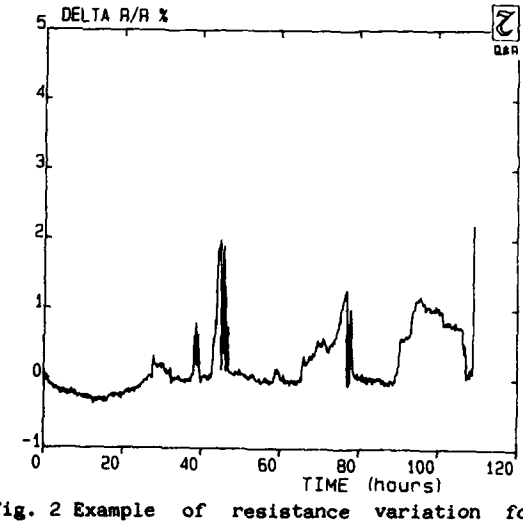


Fig. 2 Example of resistance variation for a sample showing multiple resistance spikes before the opening at T=223°C.

On analyzing these plots, the resistance variations seem to occur randomly during life tests. All the stripes exhibited random resist

TABLE I - Characteristics of the stripes

			Mean	Stand. Dev.
Temperature coefficient (with reference to 25°C)	α	[°C^{-1}]	$3.969.10^{-3}$	$0.087.10^{-3}$
Resistance at 25°C	R	[Ω]	3.641	0.156
Thermal resistance	R_{th}	[°C mW^{-1}]	1.661	0.085
Joule heating (at 176°C and I=50 mA)	ΔT	[°C]	25.7	1.8
Sample cross section	S	[μ^2]	1.57	0.066
Current density	J	[A.cm^{-2}]	$3.186.10^6$	$0.138.10^6$

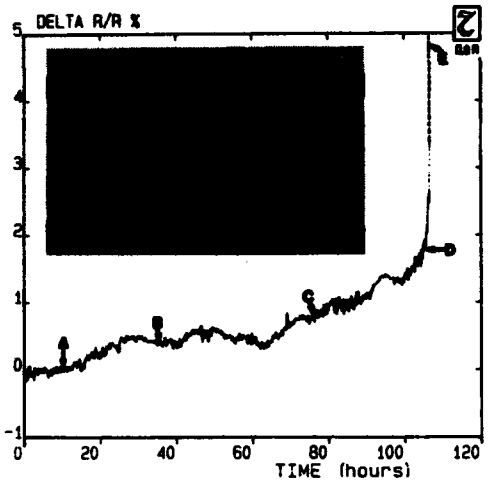


Fig. 3 Example of resistance variation for a sample showing a progressive resistance increase. The inset displays a SEM picture of the stripe after opening. Letters link the resistance variation model to the actual damage.

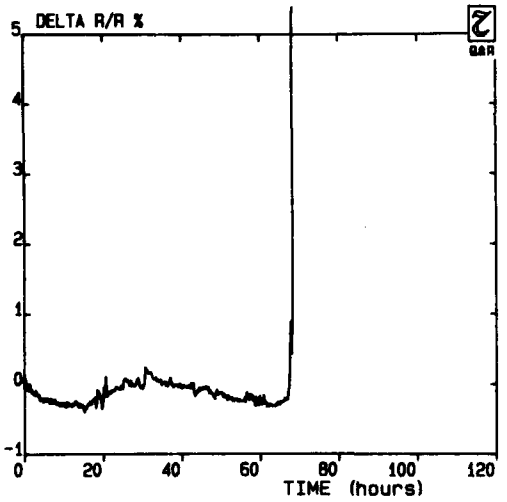


Fig. 4 Example of resistance variation for a sample showing abrupt failure at T=223°C

ance spikes of varying duration and amplitude and the total resistance variation was generally less than 6%.

Four common behaviours were roughly identified, whose frequency of occurrence is indicated in parentheses: a) the mean resistance is roughly constant with superimposed multiple resistance spikes, before the final opening (17.7%), as in fig. 2; b) the mean resistance increases with superimposed multiple resistance spikes (40.3%); c) the resistance increases with no or very limited spikes, suddenly the curve becomes very sharp and the stripe fails (24.2%), as in fig. 3; d) the means resistance is nearly constant with no or very limited spikes, and then fails abruptly (17.8%), as in fig. 4.

5. DISCUSSION

Our results can be explained by taking into account the structure of the stripes. The high value of the TTF dispersion (σ=1.3) was not unexpected, owing to the influence of structural and dimensional fluctuations in so narrow stripes.

We tried to correlate the evaluated TTF to the resistance variations for each piece, but we did not find any correlation between the TTF and the resistance variation or the occurrence of the first spike.

The spikes can be explained as an aggregation and vanishing of voids, which temporarily varies the cross section of a stripe at some point [7]. In any case the final failure seems to occur in a completely random way, with no correlation with the spikes.

Our results on the resistance behaviour, partly disagree with those of some researchers [3-6] who found a regular increase in the stripe resistance before failures. These differences are explained by the differences in the dimensions and structure of the tested stripes. In fact our stripes have a bamboo structure as can be seen in fig. 5a. The stressed stripes presents only few damaged areas, where EM starts (see fig. 5b). Frequently the shape of the voids was triangular (fig. 5c-d) and they grew by starting from the bottom of the stripe, without reaching the top surface. Owing to the very simple appearance of the damage, we tried to model the percentage $\Delta R/R$ variation with the void growth and we postulated the growth of triangular - or trapezium - shaped prisms; the model well described the regular $\Delta R/R$ increase of fig. 3.

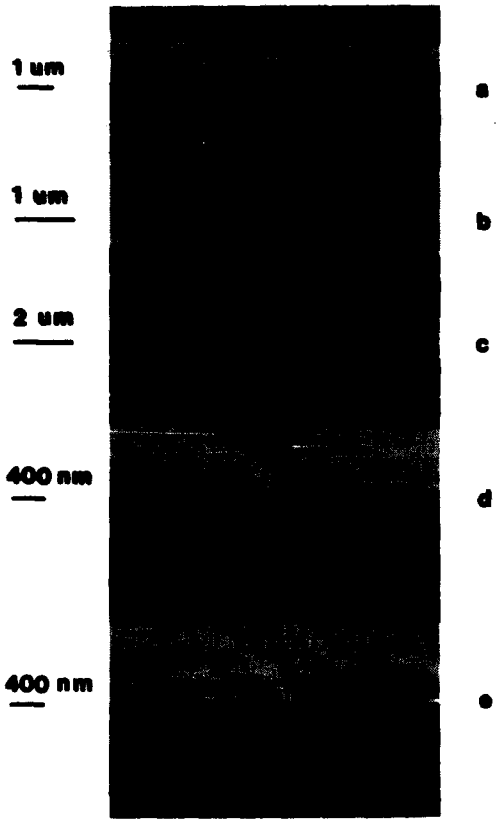


Fig. 5 SEM pictures of stressed stripes. The arrows in picture a) indicate the grain boundaries.

Some stripes exhibited only very limited damages, like the very narrow cut shown in fig.5e, about 70 nm wide, as can be roughly estimated from the SEM picture. This kind of failures are found when the resistance does not increase during the test but the opening occurs abruptly. This last failure mechanism cannot be described by the resistance variation, and severely limits the effectiveness of quick techniques,like TRACE [4], for the reliability evalutation of very narrow stripes.

6. CONCLUSION

The electromigration tests performed on VLSI Al(Si) stripes enables us to ascertain that:

i) the estimated standard deviations of the time to failure distribution are higher than that of wider striper, in spite of the corrections made, so only due to the influence of structural and dimensional fluctuations in so narrow stripes;

ii) the increase of stripe resistance before failure does not occur for all the stripes and so cannot be correlated with the end of life;

iii) several resistance spikes are observed before the complete opening of the stripes that are enhanced by the small dimensions of the stripes and can be related to void formation and vanishing; however for this phenomenon, to, no correlation with the end of life can be established;

iv) a simple degradation model was developed that justifies the observed resistance variations.

ACKNOWLEDGEMENTS

The authors acknowledge the contribution of R. Bottini and M. Vanzi for the failure analyses, of C. Corradini for collecting the data and of C. Caprile for sample preparation.

REFERENCES

[1] Ho, P.S., IEEE an. Proc. Reliab. Phys 20 (1982), 288.

[2] Black, J.R., IEEE An. Proc. Reliab. Phys. 6 (1967), 148.

[3] Rosenberg, R. and Berenbaum, L., Appl. Phys. Lett. 12 (1968) 201.

[4] Hummel, R.E., Dehoff, R.T. and Geir, M.J., J. Phys. Chem. 37 (1976), 73.

[5] Rodbell, K.P. and Shatynsky, S.R., Thin Solid Films, 108 (1983) 95.

[6] Pasco, R.W. and Schwarz, J.A., Solid State Electronics, 26 (1983) 445.

[7] La Combe, D.J. and Parks, IEEE an. Proc. Reliab. Phys. 23 (1985) 74.

[8] Specchiulli, G., Milani, F., Scaccabarozzi, M., in 3rd Int. Conf. on Quality in Electronic Components (Bordeaux, 1987) pp. 92-100.

[9] Bobbio, A., Ferro, A. and Saracco, O., IEEE Trans. on Reliability R-23 (1974) 194

Electromigration Control : the Accelerated BEM (Breakdown Energy of Metals) Test

L. Bacci, C. Caprile, G. DeSanti
SGS Microelettronica,Central R&D, Via C.Olivetti 2
20041 Agrate Brianza (MI)-Italy

Abstract

We have used a current ramp based method (BEM) to study the electromigration response of metal lines prepared under different experimental conditions. The samples were all Al1%Si; BEM discriminated among the metallizations prepared with different sputtering systems: a better behavior was obtained for films prepared in high vacuum and with a fast deposition rate. Furthermore we have characterizad the use of both DC and RF bias for Al planarization. The results show a loss of reliability due to the bias induced crystallographic disorder of the Al structure.

1 Introduction

In the last few years the device reliability concern has strongly increased due to the stringent requirements of VLSI technologies. One of the most critical steps in the device fabrication is the final metallization whose degrade is most often due to aluminum electromigration. Among the many fast techniques proposed for electromigration control [1,2,3] the BEM (Breakdown Energy of Metals) [4,5] has the advantage to suit both to the production environment (allowing a fast and automatized check of the production lots) and to the R&D requirement to attain an evaluation of new materials and processes.

We have applied the BEM to the evaluation of both sputtering equipment and new metallization processes. In the first case the results provided a correlation between the machine characteristics and the electromigration yield of the films; in the second, the BEM results showed a strong influence of the deposition parameteres on the Al reliability. In both cases it turned out that the microstructural feature which is most sensitive to the metal film electromigration performance is the crystallographic ratio R[111]/[200] between the X ray diffracted peaks intensity.

The structural order of the film in fact, preventing the voids agglomeration that takes place at the stripe weakest points, results in a higher electromigration resistance of the metal line.

The test structures employed for the experiment consisted of straight metal lines, of different widths and lengths, laying on a flat oxide substrate. We intentionally avoided any process or architecture induced thinning of the metal section so that the electromigration performance of the interconnects could be directely related to the intrinsic properties of the material.

2 The Method

The BEM techique, introduced by Hong et al.[4], allows a wafer level electromigration testing by means of a current ramp forced in the test line; at every step the voltage drop is monitored and the temperature is calculated. When the stripe opens up the total energy transferred to the line is calculated by means of the :

$$E_F(mJ/cm) = \frac{1}{L}\sum_j I_j^2 R_j exp(-Ea/kT_j)\Delta t$$

The figure of merit of the technique is the MEF: Median Energy to Fail (mJ/cm) defined as the 50% point of the lognormal distribution of the failure energies measured on a wafer (fig.1).

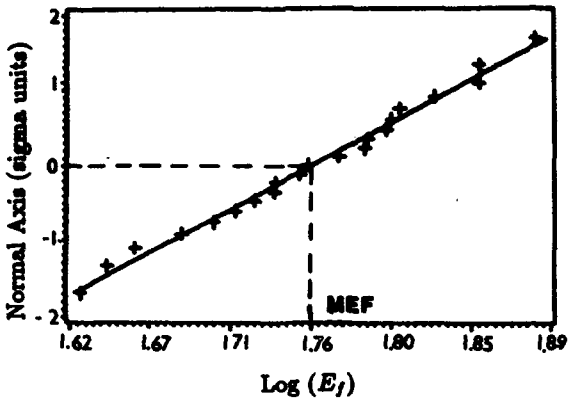


fig.1 A typical lognormal distribution of failure energies on wafer; the normal axis is centered on the distribution 50% point, +1 and -1 points correspond to +1 and -1 sigma of the distribution. The MEF value is also shown.

The stripe temperature is a function of both the current density and the heat dissipated by the metal toward the substrate. Since the temperature is an accelerating parameter for electromigration, the BEM results need to be normalized to the thickness of the oxide underlaying the metal, which constitutes the major heat sink for the stressed stripes.
We have therefore prepared a set of wafers with oxide layers of different thicknesses, from 0.1 μm to 1.5 μm, and performed the BEM test on metal films deposited on the different substrates. The test was performed on both passivated and not passivated samples.
The results are presented in fig.2 and fig.3; the MEF values are higher when the metal sits on thinner oxide layers, where the heat dissipation through the dielectric is more efficient; at thicker oxide layers the faster temperature increase results in a stronger acceleration of the phenomenon and therefore in smaller MEF values. It is important to observe that even when the rate of heat dissipation is very poor the failure mechanism is still electromigration; this is confirmed both by the monotonic behavior of the curves of fig.2 and fig.3 and by SEM observations of the stressed stripes. As expected the effect of a passivating layer is to increase the electromigration performance of the metal and to decrease the dispersion of the failure energies on wafer.

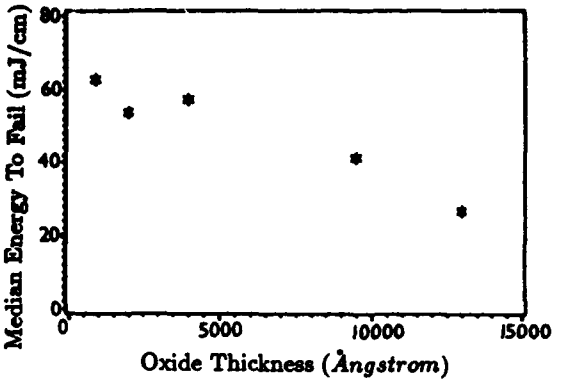


fig. 2 MEF values for a set of samples with different thickness of the substrate oxide; not passivated samples.

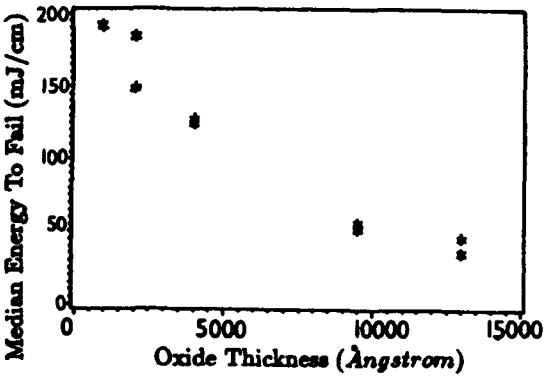


fig. 3 MEF values for a set of samples with different thickness of the substrate oxide; passivated samples.

3 Experimental

The tests were all performed on 4" Silicon wafers metallized with 1 μm thick Al1%Si films, on four terminals stripes of dimensions 4 μm - 320 μm and 8 μm - 640 μm. The substrates were oxidized prior the metal deposition. In all cases the passivation was 7000Åof SiO_2 phosphorous doped at $4-5\%$.
The wafers for equipment characterization were metallized on four sputtering systems whose main features in terms of vacuum configuration and of deposition rate are described in table I. In this respect the equipment are rapresentative of the evolution in the technological approach to VLSI metallization that has taken place in the last years. For this experiment the wafers were oxidized with a 1.7 μm thick SiO_2 layer.

Equip.	MEF (mJ/cm)	R [(111)/(200)]	Chamber pump.syst.	Loading system	Pre-sputter. pressure	Deposition rate
A	310	9.5	diffusion pump	no load lock	$3*10^{-6}$torr	250 A/min
B	325 338 383	871	cryo pump	load lock (mechan. pumped)	$8*10^{-7}$torr	3500A/min
C	484 538 629	(111) only	cryo pump	load lock (cryo pumped)	$2*10^{-7}$torr	1.0um/min
D	531 541 501	(111) only	cryo pump	load lock (cryo pumped)	$1*10^{-7}$torr	1.7um/min

Table 1.

The realization of multiple level metallizations is one of the key points for increasing the degree of VLSI device integration. Multilevel metal devices are extremely critical in terms of surface topography; a poor step coverage of the metal layers, especially in contacts and vias with high aspect ratio, can therefore result in a reliability concern for open circuit failures. Aluminum film planarization is one of the proposed solutions, since it provides adequate sidewall coverage and complete via filling.

Between the prospected methods for increasing the metal redistribution on wafer are the high temperature deposition and the application of a biasing voltage to the substrate: the first enhances the mobility of the deposited Al atoms, the other, if the bias is high enough with respect to target power, increases the resputtering rate.

We have investigated the effect of substrate biasing on Al film reliability. The samples were prepared using equipment C and D (table I) applying both DC and RF bias to the substrate. The underlaying oxide was 7000Åfor all these samples.

Film crystalline structure was investigated using X-ray diffractometry. The significant parameter for evaluating the degree of cristalline order has been taken to be the ratio between the height of the diffraction peaks corresponding to the crystallographic planes (111) and (200) of aluminum.

4 Results and Discussion

In table I we present the results of the BEM test performed on Al1%Si films prepared with the four deposition systems. MEF values are normalized to the initial stripe resistance (R_0) to compensate dimensional variations due to photolitographic definition.

Our results indicate a strong correlation between the purity level of sputtering atmosphere and the good quality of the deposited films, both in terms of the crystallographic growth and of the electromigration performances. First of all an efficiet pumping system, able to lead to pressure in the $10^{-7}torr$ range in the main chamber, is necessary to guarantee a good Al film quality (B,C and D). Pumping efficiency can be further increased introducing a transfer chamber (load-lock) to load the wafers in the sputtering area. It is also important to increase the deposition rate (C and D) . All these features are related to the reduction of gas trapping in the film during deposition, which results in a more ordered crystallographic growth of the Al grains and finally in a good electromigration behaviour.

Sample	Equipment	MEF (mJ/cm)	R [(111)/(200)]	Process parameters Temperature	Bias
1	D	80	(111) only	150 C	NO
2	D	9	20	250 C	- 215 V (RF)
3	D	8	180	150 C	- 250 V (DC)
4	C	81	(111) only	25 C	NO
5	C	75	2000	25 C	- 150 V (DC)

Table 2.

To analyze bias effects, different kinds of test samples were prepared, as shown in table II, where are reported the MEF results with the corresponding XRD data. In this case differences in R_0 are mainly due to the different deposition processes and therefore MEF values were not normalized. Samples 1 and 4 were prepared with the standard process parameters used for equipments C and D. The corresponding MEF values, the highest in this test, are typical for these sputtering systems. Crystallographic structure, showing a complete orientation along the (111) planes, indicates a well ordered growth of the films during deposition. Conversely films prepared with an high degree of substrate biasing (no matter if RF or DC bias,

samples 2 and 3) show a dramatic decrease in the MEF values with the corresponding weak (111) orientation. The low value of DC bias employed for sample 5 preparation results in a limited Al atoms resputtering and therefore in a still high MEF value. The crystal growth is not significantly modified, but no improvement in step coverage is observed.

5 Conclusions

In this work we have shown that BEM provides a fast and versatile technique for metallizations reliability testing. Our approach has been mainly in the determination of the deposition conditions influence on the reliability of the films, both in terms of equipment and of process parameters. The results show a definite relation between the electromigration resistence of Al films and their cristallographic quality; the reduction of mass flux divergence points results in a prolonged life for the conductor. The BEM results indicate that the realization of a reliable metal structure is connected on one side to the sputtering environment (good vacuum levels, fast sputtering rate) and on the other to process conditions that favor the ordered growth of the films.

For metal planarization our results point to the risks involved in the use of a substrate bias either DC or RF. If the biasing voltage is effective for a step coverage improvement it has detrimental effects on reliability, since it modifies the crystallographic growth (increase of the resputtering rate and contaminants inclusion). On the other hand if the biasing voltage doesn' t influence the Al microstructure it is not even effective in in improving the film planarization. The action of a high temperature deposition on both planarization and reliability is still under investigation.

Other applications of the technique are in progress. In the production environment BEM is used to continuosly check the quality of the outgoing lots by means of a test stripe inserted in the devices scribing line. Evaluation of new metallization alloys is a future application, once the activation energy for electromigration in the materials under study will be known.

6 Acknowledgments

This work has been partially supported by the EEC inside the ESPRIT Project 554 SPECTRE. We wish to thank F.Riva for the experimental setup, G.Queirolo for the XRD measurements and M.Mora for supplying some of the samples.

7 References

1) R.W.Pasco, J.A.Schwartz Solid State Electron. 26, p.445 (1983)
2)T.M.Chen,T.P.Djeu IRPS 1985 Proceeding, p.87
3)B.J.Root,T.Turner IRPS 1985 Proceedings, p.100
4)C.C.Hong,D.L.Crook IRPS 1985 Proceedings, p.108
5)L.Yau,C.Hong,D.L.Crook IRPS 1985, p.115

ELECTRICAL AND STRUCTURAL CHARACTERIZATION OF ELECTROMIGRATION IN Al-Si/Ti MULTILAYER INTERCONNECTS

M.FINETTI, A.SCORZONI, A.ARMIGLIATO, A.GARULLI

CNR - Istituto LAMEL, Via Castagnoli 1, 40126 Bologna, Italy

I.SUNI

Semiconductor Laboratory, Technical Research Centre of Finland, Otakaari 7B, 02150 Espoo, Finland

Abstract

In this work we investigate the correlation between the resistance behaviour and the mass transport in narrow line interconnects with a laminated structure incorporating multiple layers of Ti in Al-Si. For comparison, homogeneous Al-Si films are also studied. The electromigration resistance is evaluated by applying a temperature-ramp resistance analysis (TRACE) at direct wafer level. The improved stability against electromigration in the laminated structures is related to the formation of a continuous barrier of the intermetallic compound $TiAl_3$, preventing voids from propagating across the film.

1. INTRODUCTION

Electromigration lifetime of Al or Al-Cu interconnections can be significantly improved by incorporating a compound of Al and a transition metal inside the metallization layer /1,2,3/.

In this paper we present the results of an experimental work performed to correlate the electromigration induced damage to the resistance changes in homogeneous Al-Si and laminated Al-Si/Ti stripes.

The test lines were investigated by applying a temperature ramp-technique (TRACE) /4/. By using this method to analyze the resistance change in a stripe under high current stress, the kinetic parameters for the early stages of electromigration can be determined.

Failure analysis and metal film characterisation were carried out by optical microscopy, SEM and TEM analyses.

2. EXPERIMENTAL

The metal films were deposited on 3 inch oxidized silicon substrates. Two different types of metallizations were prepared, using magnetron sputtering. The first metallization consisted of four Al-Si layers, 1100 Å thick each, alternating with three 200 Å thick Ti films deposited during the same pumpdown. In the following this will be referred to as multilayer structure. For a direct comparison of the electromigration behaviour in unlaminated structures, Al-Si films of 5000 Å thickness were also prepared. After sputtering, different types of electromigration test vehicles (see ref.5) were delineated by standard photolithography and chemical wet etching. The results reported in this paper are relative to linewidths ranging between 4 and 5 μm. The samples were then annealed in an argon gas flow at 450°C for 30 min, in order to form the intermetallic $TiAl_3$.

The electromigration measurements were carried out at wafer level using a manual probe station with a temperature controlled hot chuck. The stripes were tested using the TRACE method /4/. In this technique, the elec-

tromigration kinetics relative to the early stages of electromigration are studied by performing resistance measurements as a function of a linearly ramping temperature. The test set-up was entirely controlled and monitored using an HP-9816 computer. The increase in the stripe temperature induced by the stress current was determined by measuring the resistance at different current densities between room temperature and 473 °K. As a result the electromigration effects are displayed as a function of the actual temperature of the stripes. Optical microscopy, SEM and cross-sectional TEM were used to investigate the structural damage caused by the current stress.

3. RESULTS AND DISCUSSION

The first experiments were performed on the Al-Si metallization. By following ref.4, the measured conductor resistance was separated into two independent additive terms, R_T and R_{EM}, representing the temperature and the electromigration component respectively. The measurements reported in this work were carried out at $J=2.8\times10^6$ and 5.5×10^6 A/cm^2 and ramping rate r=0.25°K/s. A typical result obtained on an unlaminated Al-Si film under current stress is shown in Fig.1, in which

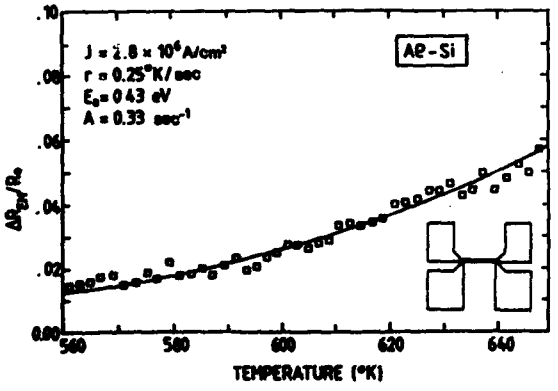


FIGURE 1

Electromigration resistance as a function of stripe temperature

ΔR_{EM} is the resistance increase induced by electromigration and R_o is the resistance at the initial ramping temperature. The data for the initial 6% relative resistance change fit a thermally activated process described as

$$\frac{\Delta R_{EM}}{\Delta t}\ \frac{1}{R_o} = A \exp\left(-\frac{E_a}{kT}\right)$$

with an activation energy E_a=0.43 eV and a pre-exponential factor A=0.33 s^{-1}. The stripes stressed at 5.5×10^6 showed similar activation energies, but in this case they failed by developing an open crack at an average temperature of about 613°K. Randomly located small hillocks appeared along the line during the test (Fig.2).

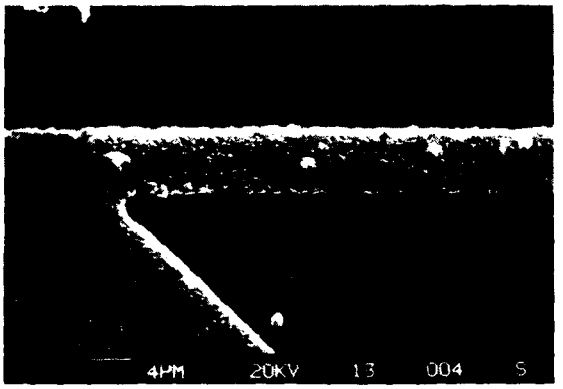


FIGURE 2

SEM micrograph of an unlaminated Al-Si structure stressed at J = 5.5×10^6 A/cm^2, showing hillocks randomly located along the line.

In order to compare the behaviour of the different metallizations under study, the multilayer structures were tested at the same current density ($J=5.5\times10^6$ A/cm^2), while for comparison with the unlaminated Al-Si structure the stripe temperature was ramped up to the failure temperature of these films (613°K). In this case, the resistance monitored as a function of the stripe temperature did not show any deviation from linearity, indicating that the electromigration component in the measured resistance is completely.

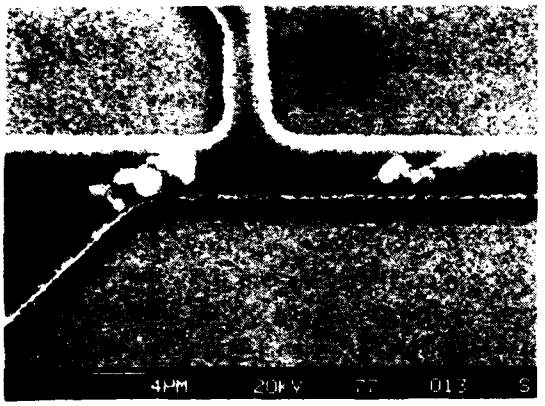


FIGURE 3

SEM image of a multilayer structure stressed at $J=5.5\times10^{6}$ A/cm^{2}, showing hillocks formation at the positive end of the line.

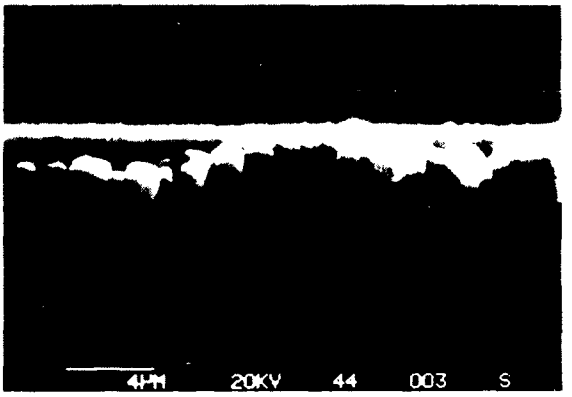


FIGURE 4

Large extent of hillocks along the positive end of the multilayered stripe.

negligible. However, SEM observations of the stressed multilayer structure revealed a limited amount of hillocks (see Fig.3). In contrast to the case of the unlaminated Al-Si the accumulation of material occurred only at the positive end of the stripe. Slight erosion of material was observed at the junction of the stripe and cathodic current terminal but it was not effectively detected in the voltage measurement since the depletion mainly occurred between the current tap and the voltage tap. In fact, increasing the stripe temperature to 663°K the resistance change (deviation from linearity) became negative /5/. This was concurrent with an increasing number of hillocks which appeared along the positive end of the line extending beyond the voltage tap, and were therefore electrically detectable (see Fig.4). The observed behaviour is in agreement with the results obtained from accelerated lifetime tests as reported in ref.2.

The extensive formation of hillocks was never found in homogeneous Al-Si films. One should note however that single layer metallizations failed by developing open cracks below the temperature at which the formation of hillocks on multilayer stripes effectively sets in.

To avoid the effect of hillocks on the voltage measurement in the multilayer structures, this was performed as shown by the inset in Fig.5. For relative resistance changes up to 5% the experimental data could be adequately described by a single activated process with $E_a = 0.54$ eV and $A = 2.7$ sec^{-1}. The low value of the activation energy indicates electromigration processes directly related to Al grain boundary diffusion or to another mechanism with a fortuitously similar activation energy. Further measurements at different ramping rates are currently under way.

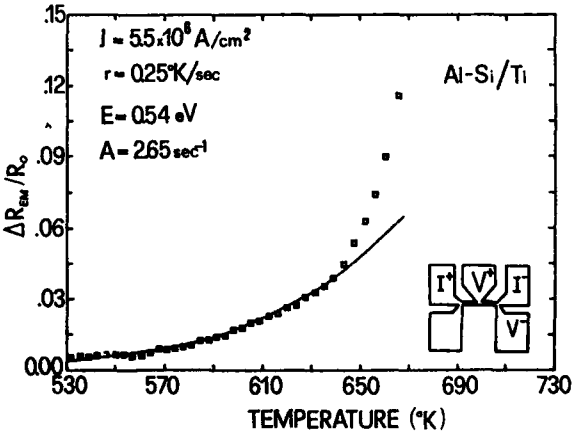


FIGURE 5

Electromigration resistance as a function of stripe temperature (see ref.5 for details).

The limited number of electromigration induced defects on the test stripe (except in the stripe ends) would suggest a more uniform microstructure in the laminated metallizations compared to the conventional Al alloy films. Other investigators have pointed out that relatively uniform columnar grains and a continuous intermetallic compound layer are fundamental in improving the electromigration resistance of laminated interconnects /1/. An appropriate TEM procedure for sample preparation has been recently set up /6,7/ to obtain longitudinal cross-sections in 5 /um

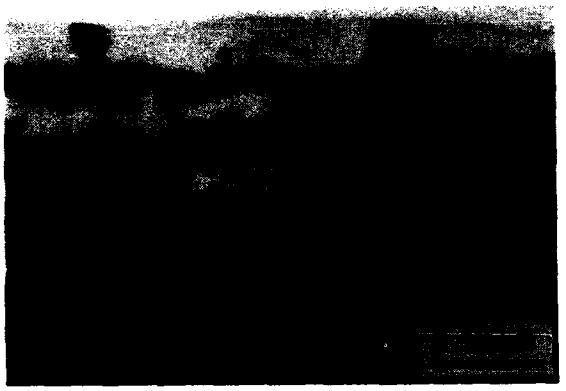


FIGURE 6

Bright-field TEM micrograph of an Al-Si/Ti multilayer structure annealed at 450°C for 30 min and stressed at $J=5.5\times10^6$ A/cm^2 up to 663°K.

wide stripes. Fig.6 shows the cross-sectional image of a multilayer Al-Si/Ti structure annealed at 450°C for 30 min and stressed at $J=5.5\times10^6$ A/cm^2 up to a stripe temperature of 663°K. A continuous, although not-uniform, layer of $TiAl_3$, as identified by electron diffraction analysis, has formed. Since the intermetallic phase prevents voids from propagating across the film and is able to carry large current densities, the barrier continuity is of primary importance. Furthermore, the different layers are not completely planar, as shown by their "wavy" morphology. Due to the difference in electrical resistivities of Al and $TiAl_3$, interface roughness could give rise to large current density variations. For thin Al overlayers this could become detrimental, whereas for thicker films the effect of interface roughness is expected to be small.

4. CONCLUSIONS

The resistance degradation caused by electromigration proceeds in two different ways in single layer Al-Si and multilayer Al-Si/Ti films. For Al-Si interconnects, stressed at 5.5×10^6 A/cm^2, the degradation starts at relatively low temperatures leading to open-crack failure at about 613°K. For multilayer interconnects the resistance change is only observed above this temperature, although the activation energies suggest similar electromigration processes. The structural damage in laminated interconnects under high current stress is initially constrained to the stripe ends where the Al migration results in erosion or accumulation of material, depending on the flux direction. The cross-sectional TEM observations confirm that a continuous barrier of the intermetallic compound has formed inside the laminated structures during the annealing stage. It is quite evidently responsible for the improved stability against electromigration failure.

REFERENCES

/1/ J.K.Howard, J.F.White and P.S.Ho, J.Appl. Phys., 49 (1978) 4083.

/2/ S.S.Iyer and C.Y.Ting, Proc.22nd Int.-Symp. on Reliability Physics, IEEE Electron Devices and Reliability Soc., Las Vegas, NY, 1984, p.273

/3/ D.S.Gardner, T.L.Michalka, K.C.Saraswat, T.W.Barbee, J.P.McVittie and J.D.Meindl, IEEE J.on Solid State Circuits, SC-20 (1985) 94

/4/ R.W.Pasco and I.A.Schwartz, Solid-State Electron., 26 (1983) 445

/5/ M.Finetti, H.Ronkainen, M.Blomberg and I.Suni, Proc.1985 MRS Meeting (Boston), R.J.Nemanich,P.S.Ho and S.S.Lau Editors, pag.812.

/6/ A.Garulli, A. Armigliato and M.Finetti (to be published)

/7/ M.Finetti, I.Suni, A.Armigliato, A.Garulli and A.Scorzoni, to be published on J. Vacuum Science and Technology, Oct.1987.

PLASMA ENHANCED CHEMICAL VAPOUR DEPOSITION OF TUNGSTEN ON SILICON AND SILICON DIOXIDE SUBSTRATES

C.M.T. Hodson, J. Wood, and M. Middleton

Department of Electronics, University of York,
Heslington, York YO1 5DD,
England.*

Tungsten films have been deposited onto silicon and silicon dioxide substrates by plasma enhanced chemical vapour deposition, using the reaction of tungsten hexafluoride and hydrogen. The deposition rate, sheet resistance, and adhesion of the films have been studied as a function of the process parameters: substrate temperature, gas flow and composition, and pressure. Adherent tungsten films with sheet resistances lower than 1 ohm per square have been produced.

1. INTRODUCTION

In recent years there has been considerable interest in the use of refractory metals, and their silicides, as metallisations for silicon ICs. These metal systems have several advantages over conventional aluminium-based metallisations for electrical contacts, interconnections and vias at the VLSI level of miniaturisation [1,2]. Tungsten has been identified as being particularly suitable: it has high conductivity, is thermally stable and relatively inert, and does not suffer electromigration problems [3].

The common techniques for the deposition of tungsten are sputtering and Chemical Vapour Deposition (CVD). The advantages of CVD processes are well-known, the most important being the ability to deposit conformally over steps in the underlying structure and the opportunity to vary the film composition as a function of depth [4,5]. Tungsten lends itself well to CVD processes by virtue of the existence of the volatile and easily reducible hexafluoride WF_6. This can be reduced to tungsten by reaction either with hydrogen or at the silicon surface. Plasma-enhanced processes offer a further benefit in that the rate of reaction can be increased significantly without raising the substrate temperature: this is the regime of plasma-controlled processes.

One of the objections to the use of PCVD is that the dependence of film growth on the process conditions is often poorly understood. The purpose of our work is to carry out a comprehensive programme of characterisation of the deposition of tungsten on silicon in a glow discharge. So far the effects on the reaction rate of gas flow and composition, substrate temperature and chamber pressure have been studied [6].

2. THE PLASMA CVD EQUIPMENT

All the deposition studies were carried out in a Plasma Technology planar diode reactor, as shown in figure 1. The chamber is pumped by

*This work has been carried out with the support of the Alvey Directorate and the Science and Engineering Research Council, U.K.

a nitrogen-ballasted Roots pump combination, with the pump speed - and hence chamber pressure - controlled by a throttle valve. The reactor features purpose-built gas handling equipment and control electronics. The power supply for the glow discharge is a 0-800 Watt, 10-120 kHz, generator. Further work to investigate the characteristics of the process in the MHz region will be carried out shortly.

The silicon substrates are placed on the earthed lower electrode, which is resistively heated, allowing the subatrate temperature to be controlled up to 400°C. This forms the basis of a set of process experiments outlined below.

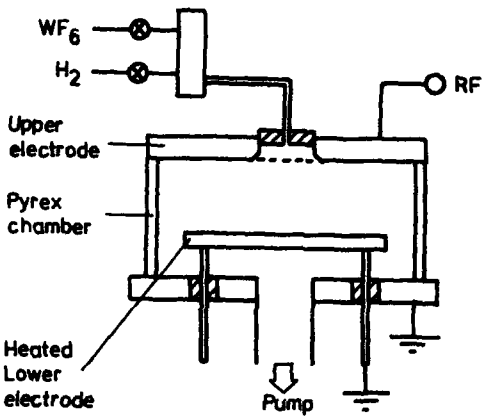


Figure 1: Schematic diagram of the plasma deposition system.

3. PROCESS CONDITIONS

A standard set of process conditions was arbitrarily chosen as a basis for comparisons between experiments. The standard process is as follows:

Chamber pressure	250 mTorr
WF_6 flow rate	10 sccm
H_2 flow rate	30 sccm
Substrate temperature	250°C
rf Power	25 W @ 100kHz

The gas flow rates are in the stoichiometric ratio for the hydrogen reduction of tungsten hexafluoride:

$$WF_6 + 3H_2 \rightarrow W + 6HF$$

The rf power corresponds to a power density of 5 $mWcm^{-2}$ in this system.

In the work described here we have varied:

(1) The ratio of WF_6:H_2 between 1:1 and 1:10;
(2) The WF_6 flow rate between 0 and 15 sccm;
(3) The substrate temperature between 150°C and 400°C;
(4) The pressure between 200 and 1000 mTorr.

The silicon and oxide coated wafers were given a vapour clean in organic solvents, followed by an RCA clean [7], which was found to improve adhesion of the films significantly, although it did not affect the properties of the films themselves.

4. RESULTS AND DISCUSSION

The quality of the deposited film surface was assessed by optical microscopy using Nomarski interference contrast. The deposition rate was determined by microgravimetric measurements, and the sheet resistance of the films was measured using a four-point probe. The sheet resistance was approximately inversely proportional to the deposition rate, confirming the general observation of Green and Levy [8] that the resistivity of tungsten is approximately constant for films greater than 50 nm thickness.

The deposited films have been in the range 50-460 nm thickness, with resistivities of between 15 and 30 microhm-cm. The lowest sheet resistances observed were around 0.6 ohms per square in films grown at 350 to 400°C; the process time was 20 minutes.

Most of the work has been carried out on silicon substrates, where the initial reaction is one of silicon reduction of the WF_6 to form an initial tungsten layer [9]. Once all the silicon sites have been used up the deposition proceeds by hydrogen reduction of WF_6.

The adhesion of the tungsten films to silicon is generally good, although the adhesion and deposition rate can be affected by deposition on the chamber walls. Films have been deposited on oxide coated silicon wafers, and also onto glass. Adhesion in both cases is worse than that to silicon at high substrate temperatures, but below 250°C they are comparable. It is hoped to produce a completely non-selective process by means of a suitable surface treatment.

4.1 Variation of the gas flow ratio

The lowest resistivity films were found to occur when the gas flow ratio was approximately stoichiometric, such that WF_6:H_2 was equal to 1:2.5. At lower hydrogen concentrations competition occurred between deposition and etching reactions due to the presence of excess free fluorine. At higher hydrogen concentrations the deposition rate was reduced, possibly due to dilution effects. The ratio of WF_6 to H_2 was fixed at 1:3 for all subsequent experiments.

4.2 Variation of the gas flow rate

The deposition rate was highest at a WF_6 flow of about 10 sccm, and the rate fell rapidly as the flow rate was reduced. This suggested that the process was limited by mass transport; that is, by the rate at which the WF_6 reached the reaction chamber. The reduction in deposition rate as the flow rate was increased above 10 sccm may be due to a reduced dwell time of active species in the chamber. The film quality and adhesion were better at the lower flow rates, where t. deposition rate was lower.

4.3 Substrate temperature

The temperature dependence of the deposition rate was of exponential form, indicating that the reaction which took place was of an heterogeneous character. The apparent activation energy of the plasma enhanced reaction was 0.23 eV per atom, compared with the published value of 0.71 eV per atom for the CVD process [10]. This suggested that the rate-limiting step was either a direct ion-surface reaction replacing the rate-limiting hydrogen adsorption step in the CVD case, or that the presence of ion bombardment was reducing the apparent activation energy of the reaction. The Arrhenius plot of the sheet resistance is shown in Figure 2.

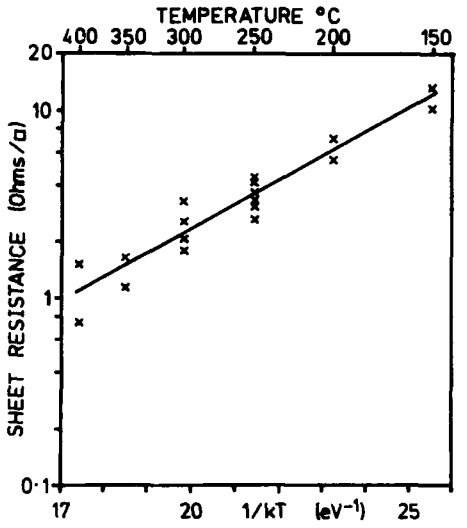


Figure 2: Arrhenius plot of the sheet resistance of the tungsten films

4.4 Chamber pressure

To investigate further the chemical behaviour of the reaction, the variation of the deposition rate with chamber pressure was studied. The gas flow rates were maintained constant, and the pressure varied by adjusting the pumping speed. Homogeneous reactions are largely insensitive to pressure variations. The reaction occurs in the gas phase and any solid products may then deposit on the wafer. In contrast, heterogeneous reactions occur by adsorption of the molecules onto the surface where they then react. Such reactions depend strongly on pressure but not on gas flow [12,13]. The initial results, displayed in Figure 3, show that above about 300 mTorr in the chamber, the deposition rate was almost independent of pressure, and may thus be classified as homogeneous. For the reaction

vessel we are using, pressures below about 200 mTorr should promote the heterogeneous character of the reaction. Work is currently under way to investigate the low pressure regime in order to characterise this deposition reaction.

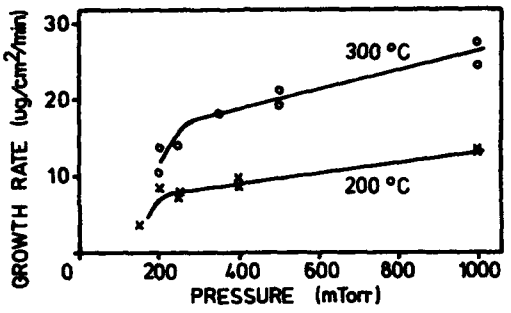


Figure 3: Dependence of growth rate on pressure

5. CONCLUSIONS

Adherent, low resistivity tungsten films have been deposited successfully onto silicon substrates using Plasma-CVD techniques. Films can be grown with sheet resistances below 1 ohm per square in 20 minutes at process temperatures of between 250 and 400°C.

Further studies of the chemical kinetics are in progress to investigate the effects of the rf discharge on the reaction rates and mechanisms.

REFERENCES

[1] Roland, J.P., Hendrikson, N.E., Kessler, D.D., Jovy, D.E., Quint, D.W., Hewlett-Packard Journal, August 1983, 30-33

[2] Vossen, J.L., J. Vac. Sci. Tech. 19 (1981) 761-765

[3] Wolf, S. and Tauber, R.N., Silicon Processing for the VLSI Era (Lattice Press, U.S.A., 1986)

[4] Hess, D.W., J. Vac. Sci. Tech. A2 (1984) 244-252

[5] Akitmoto, K. and Watanabe, K., Appl. Phys. Lett. 41 (1982) 75-78

[6] Hodson, C.M.T. and Wood, J., Elec. Letts., in print.

[7] Kern, W. and Puotinen, D.A., RCA Review (1970) 187-206

[8] Green, M.L. and Levy, R.A., J. Electrochem. Soc. 132 (1985) 1243-1250.

[9] Pauleau, Y. and Lami, Ph., J. Electrochem. Soc. 132 (1985) 2779-2784

[10] Broadbent, E.K. and Ramiller, C.L., J. Electrochem. Soc. 131 (1984) 1427-1433.

[11] Wilkes, J.G., Solid State Devices (1985) 169-182

[12] Wilkes, J.G., J. Cryst. Growth 70 (1984) 271-279

PLASMA ANODIZATION OF SILICIDES

B. PELLOIE*, J. PERRIERE*, J.P. ENARD*, A. LAURENT*, I. MONTERO+, A. CLIMENT+, R. PEREZ+, J.M. MARTINEZ-DUART+, R. NIPOTI$ and S. GUERRI$

* Groupe de Physique des Solides de l'E.N.S., Université Paris VII, Tour 23, 2, place Jussieu, 75251 Paris Cedex 05, FRANCE.
+ Inst. Ciencia de Materiales, CSIC, and Dpto. Fisica Aplicada, C-12, Universidad Autonoma, Cantoblanco, 28049 Madrid, SPAIN.
$ C.N.R. Istituto LAMEL, Via Castagnoli, 1, I-40126 Bologna, ITALY.

We summarize here, our present understanding on the mechanisms of formation of oxide films grown on silicides by room temperature plasma anodization. The studies of the oxygen ions and cations transport processes during the plasma oxide growth on silicides of various refractory metals (Zr, Hf or Ta) and rare earths (Gd, Sm) were carried out using a combination of nuclear reaction analysis, Rutherford backscattering spectrometry and ^{18}O isotopic tracing experiments. To explain the results on ionic migration, we propose a model based on a "place exchange mechanism" in which molecular entities exchange their position during oxide growth.

1. INTRODUCTION

Due to their low resistivity and high chemical stability, silicides are finding increasing applications as gate or interconnection material in the very large scale integrated devices. Moreover, the possibility to produce on silicides a protective coating (electrically insulating oxide) offers inherent advantages in multilevel metallization schemes. To successfully utilize silicides in device fabrication, a good understanding of their oxidation behaviour as well as of the properties of the grown oxides is required. As a result, a good deal of work on the thermal oxidation of silicides has been done over the past years [1,2].

An interesting feature of the thermal oxidation of most silicides is that the growing oxide appears to be pure SiO_2, while the integrity of the silicide is preserved. This may be related, in part, to the oxygen transport through the oxide and to the rapid migration of Si and metal atoms through the silicide.

We have studied the possibility of growing oxide films on silicides by a low temperature and dry process : anodization in an oxygen plasma [3,4] . Thus, we have shown that thick oxide films can be grown on refractory metal (Zr, Hf or Ta) silicides, or on rare earth (Gd, Sm) silicides at room T. Use of the term "plasma anodization" indicates that the potential at the surface of the sample (V_s) is determined by the plasma parameters, and that the substrate to be oxidized is positively biased (V_b) with respect to V_s. The resultant mean electric field E across the oxide of thickness d is defined by :

$$E = (V_b - V_s) / d$$

This field, which is in the range 10^6 to 10^7 V/cm is the driving force for the migration of ions through the oxide leading to the growth of thicker films. One important difference from the thermal oxidation of silicides is that the plasma anodization leads to an oxide formation which does not affect in any way the silicon-silicide interface since, because of the low T of formation, any atomic transport through the silicide can be ignored.

In this paper, we present the results of our study on the oxygen ions and cations transport processes during plasma anodization of silicides. The case of silicides is by itself very attractive, since the presence of two distinct cationic species offers an alternative to isotopic tracing of cations. In this work, the combined use of isotopic tracing and transport number measurements leads to information on oxide growth mechanisms which are not compatible with the clas-

sical treatment based on point defect migration. Hence, we propose to describe the ionic movement under high field via a place exchange mechanism between molecular groups.

2. EXPERIMENTAL

The silicides were obtained either by heat treatments of thin deposited films (Zr, Hf, Gd or Sm) onto Si, or by co-sputtering form Ta and Si targets on Si substrates followed by a vacuum annealing. The plasma anodizations were carried out in the multipolar plasma set-up previously described [3,5] . In this system, the magnetic confinment of the primary electrons allows the production of dense plasmas at low pressure (5.10^{-4} Torr).

Rutherford backscattering spectrometry (RBS) was used to determine the stoichiometry of the as-formed silicides and plasma grown oxides as well as the depth distribution of the cations after the plasma anodization. The overall ^{16}O and ^{18}O contents of the oxide films were measured using respectively the $^{16}O(d,p)^{17}O^*$ and $^{18}O(p,\alpha)^{15}N$ nuclear reactions [6] . The ^{18}O depth profiles in the oxide films were obtained by the analysis of the excitation curves of the $^{18}O(p,\alpha)^{15}N$ nuclear reaction near the 2 KeV wide 629 KeV resonance[7] .

3. RESULTS AND INTERPRETATIONS

Fig. 1 represents RBS spectra recorded for Sm silicides at various steps of the plasma anodization. These results are characteristic of the behaviour of silicides since similar spectra were obtained on Zr, Hf, Ta or Gd silicides. In Fig. 1, the shape of the Sm and Si contributions indicates the formation of a mixed oxide, the thickness of which increases with anodization potential. A surface peak is observed in the Sm part and it indicates an increase in the Sm oxide concentration in the near surface region, associated with a corresponding SiO_2 depletion. In the bulk of the oxide mixture, the ratio of Sm to Si oxide concentration appears fairly constant. The Sm surface peak shows an evolution during oxide growth : up to 75 volts, the peak area increases, and then, with further potential increase it does not show any change while the overall oxide thickness always increases.

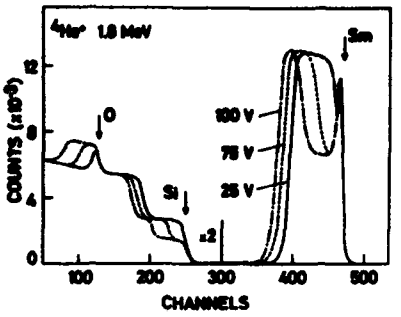


FIGURE 1

RBS spectra for $SmSi_2$ anodized in plasma at various potentials.

This behaviour, observed whatever the silicides, has been correlated to the variation of the mean electric field E during plasma anodization. E first decreases sharply with oxide thickness, then more slowly tending towards a limit value. Thus, a general description can be given: for small thicknesses, E is high (> 5.10^6 V/cm) and the surface of oxide shows an enrichment in M oxide (M for Zr, Hf, Ta, Gd or Sm) associated with a corresponding SiO_2 decrease, the reverse being observed near the substrate ; as the thickness increases, the M oxide surface enrichment increases until the value of the field has decreased to below 5 10^6 V/cm, where upon the surface enrichment in M oxide stops, while in the bulk, the oxide mixture tends towards the stoichiometry deduced from the silicide composition.

These results show that cations (M and/or Si) take part in oxide growth during plasma anodization of silicides. In fact, if we assume that the growth is only due to the movement of oxygen ions, the ratio of M to Si oxide concentration must be constant through the film, since oxygen movement cannot induce a mixing of cations. Moreover, the diffusion of Si atoms at room temperature, the migration of Si cations

against the high electric field and Si surface losses are all highly unlikely. So the surface enrichment in M oxide evidences the fact that M cations take part in the transport events. The correlation between the evolution of the oxide depth distribution and the mean electric field shows that the higher the field, the higher the enrichment of M oxide. This means that the movements are assisted by the electric field and only occur during oxide growth, since experiments carried out by biasing the sample in Ar plasma (without oxide growth) show that the oxide mixture remains unchanged.

This conclusion (movements of M cations during oxide growth) is in contradiction with the results of the measurements of the transport number of cations t^+ during plasma anodization [8]. In fact t^+, i.e. the relative fraction of oxide formed by cation migration has been investigated in oxide grown on $ZrSi_2$ using Xe atom markers. RBS was used to determine the possible change in Xe marker position related to the change in oxide thickness. For $ZrSi_2$ plasma anodization [8], the position of these markers remains unchanged during oxide growth yielding $t^+ \sim 0$. Thus, this marker experiment indicates that the oxygen ion is the sole moving species during plasma oxide growth. This leads to the following question : how is it possible to have simultaneously only oxygen ion migration and a M oxide surface enrichment suggesting M ion movements during oxide growth ?

In order to explain how oxygen movements can induce a segregation in cations, we propose an interpretation which is based on a place exchange similar in some respects to that suggested by Fromhold [9], but implying M and Si molecular groups as cation like and oxygen like complexes. We assume that a M molecular group (M_aO_x) positively charged exchanges its position with a Si molecular group (Si_aO_y) negatively charged. The high field will favour the preferential movement of M groups in the field direction i.e. towards the oxide plasma interface. As a result, the M oxide concentration will increase (with respect to Si oxide) in the near surface region.

This place exchange mechanism can be seen as the reversal of an electric dipole $(M_aO_x)^+ - (Si_aO_y)^-$ under the applied electric field, and the propagation of such an elemental event across the oxide mixture leads to the transport of charge (ionic current) from an interface to the other and to the oxide growth. In this description, there are no net cation movements since M and Si exchange their places, and there will not be oxide growth at the outer surface due to cationic migration, i.e. $t^+ \sim 0$. In fact, this place exchange mechanism leads to the equivalent movement of oxygen ion (O_{y-x}) towards the oxide-substrate interface, and this will contribute to oxide thickening at this interface with respect to the Xe markers, i.e., oxide growth only occurs by oxygen movements.

This mechanism is therefore a description of an oxygen ion movement and a consequence can be inferred. It involves the simultaneous movement under the applied field of neighbouring molecular groups, i.e. the exchange distance will not be very much higher than the interatomic distance. As a result, this should imply the conservation of the order of the oxygen atoms during an isotopic tracer experiment $^{18}O/^{16}O$.

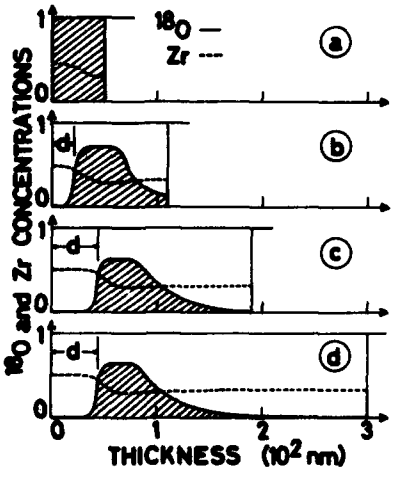


FIGURE 2

^{18}O and Zr oxide depth profiles in plasma oxide grown on $ZrSi_2$.

To check this point, such experiments were carried out and Fig. 2 represents the ^{18}O depth profiles in plasma grown oxides on Zr silicides. After the ^{18}O oxide formation (Fig. 2a), the ^{16}O plasma anodization (Fig. 2b to 2d) leads to a partial inversion of the order of oxygen atoms since ^{16}O atoms are found at the oxide-substrate interface. However, part of the ^{16}O atoms, coming from the plasma, preserves their order with respect to the ^{18}O atoms, already fixed, since a pure ^{16}O oxide layer is formed at the outer surface. The thickness of this ^{16}O surface layer increases with the overall oxide thickness up to Fig. 2c for which, the mean electric field is equal to about 5.10^6 V/cm. With further oxide growth (Fig. 2d) the mean field decreases and the thickness of the pure ^{16}O oxide surface layer remains constant. At the same time, the Zr oxide concentration, also presented in Fig. 2, which increases in the surface region up to Fig. 2c, remains constant in this region for further growth (Fig. 2d). Thus a simple analysis of Fig. 2 indicates that the pure ^{16}O oxide layer corresponds to the surface region of high Zr oxide concentration. This is evidence for a correlation between Zr oxide surface enrichment and oxygen ion migration with conservation of the order of oxygen atoms, which leads to the ^{16}O oxide layer formation.

We thus conclude that a correlation exists between the Zr cation and oxygen ion movements. Such a correlation is expected in the frame work of our assumption since we have predicted that the order of oxygen atoms will be preserved by the oxygen migration associated with the place exchange mechanism which implies cations movements.

4. CONCLUSIONS

Classically, the ionic transport under high electric field during oxide growth is explained by the thermally activated motion over discrete energy barriers of point defects like vacancies or interstitials. The results presented in this paper on the oxygen ion and cation transport processes during plasma anodization of silicides are not compatible with the classical treatment based on point defect migration. To explain these results, we propose to describe oxygen ion migration via place exchange between molecular groups, the elemental transport event involving the simultaneous movement of cations and oxygen ions.

However, such a description gives rise to numerous questions on the existence of such molecular groups, and on the ability of the oxide to support such an important disturbance during the place exchange itself. In fact, in the process the oxide network can be locally destroyed during the movement, or it may only lose its rigidity during the charge transport event. Further work is needed to give a detailed microscopic description of the place exchange mechanism.

ACKNOWLEDGEMENTS

This work was supported by the Centre National de la Recherche Scientifique (Greco 86).

REFERENCES

[1] Bartur, M. and Nicolet, M.A., J. Electrochem. Soc. 131 (1984) 371.

[2] Lie, L.N., Tiller, W.A. and Saraswat, K.G., J. Appl. Phys. 56 (1984) 2127.

[3] Perrière, J., Siejka, J., Laurent, A., Enard J.P. and d'Heurle, F., Mat. Res. Soc. Proc. 38 (1984) 443.

[4] Perrière, J., Siejka, J., Climent, A., Navarro, E. and Martinez-Duart, J.M., J. Appl. Phys. 61 (1987) 2656.

[5] Gourrier, S., Dimitriou, P., Theeten, J.B., Perrière, J., Siejka, J. and Croset, M., Appl. Phys. Lett. 38 (1981) 33.

[6] Amsel, G., J. Radio.Chem. 17 (1973) 15.

[7] Maurel, B., Thesis, Univ. Paris VII (1980).

[8] Perrière, J., Siejka, J., Rémilli, N., Laurent, A., Straboni, A. and Vuillermoz, B., J. Appl. Phys. 59 (1986) 2752.

[9] Fromhold, A.T., J. Electrochem. Soc. 127 (1980) 411.

A NOVEL PROCESS FOR SILICIDE FORMATION USING DYNAMIC RECOIL MIXING

L.I. Haworth[1)], A.E. Hill[2)], R. Holwill[1)], R. Pilkington[2)] and J.M. Robertson[1)]

1) Edinburgh Microfabrication Facility
Department of Electrical Engineering
University of Edinburgh
Edinburgh EH9 3JL, UK

2) Department of Electronic & Electrical Engineering
University of Salford
Salford M5 4WT, UK

INTRODUCTION

The limitations of doped polycrystalline silicon and aluminium as gate and interconnect materials are well known. The relatively low resistivity of metal silicides, together with their oxidisability has made them attractive candidates for replacing polycrystalline silicon gates and interconnects. At present silicide formation is achieved by metal deposition or co-deposition of metal and silicon followed by a high temperature anneal. LPCVD processing has not generally been successful for metal silicides. The dynamic recoil mixing technique described here utilises a single ion beam both to deposit metal and to mix the metal and silicon layers, and provides a low temperature route to silicide formation.

ION BEAM MIXING

A number of workers such as Tsai et al [1] and Tsaur et al [2] have used ion beam mixing to promote substantial atomic mixing at the metal-silicon interface, the projected range of the incident ions must be sufficiently large for a substantial proportion to cross the interface. Consequently, beam energies in excess of 100 keV are usually required for the commonly used species such as As^+, P^+ or Ar^+ resulting in a high level of damage which must then be annealed. The advantage of low temperature processing is therefore lost. By maintaining a thickness of metal on the substrate surface which is less than that required to form the complete silicide layer, the beam energy can be reduced by an order of magnitude to 10-40 keV, resulting in a much lower level of damage.

DYNAMIC RECOIL MIXING

The original dynamic recoil mixing (DRM) technique has been described at length elsewhere [3], [4] and is illustrated in figure 1. A 10 keV ion source is used to bombard a surface film. If the energy of the bombarding beam is adjusted so that the maximum in deposited energy is near to the metal-silicon interface considerable atomic mixing will occur. However, the bombarding beam will also back sputter the metal film which is renewed by using a second ion source (1-2 keV) to simultaneously sputter from a metal target. Adjustment of the relative fluxes of the two ion beams results in a dynamic equilibrium such that the metal layer thickness remains constant. Under such conditions significant recoil mixing occurs. Using this technique, 10 nm films of cobalt have been recoil mixed with 400 nm polycrystalline silicon layers grown on silicon dioxide [5]. Following an electron beam anneal, sheet resistivities less than 7Ω/sq were measured for the recoil mixed films. This value is 16% lower than those measured for silicide layers formed by metal deposition and given an identical electron beam anneal. Rutherford Backscattering Spectrometry (RBS) showed that the uniformity

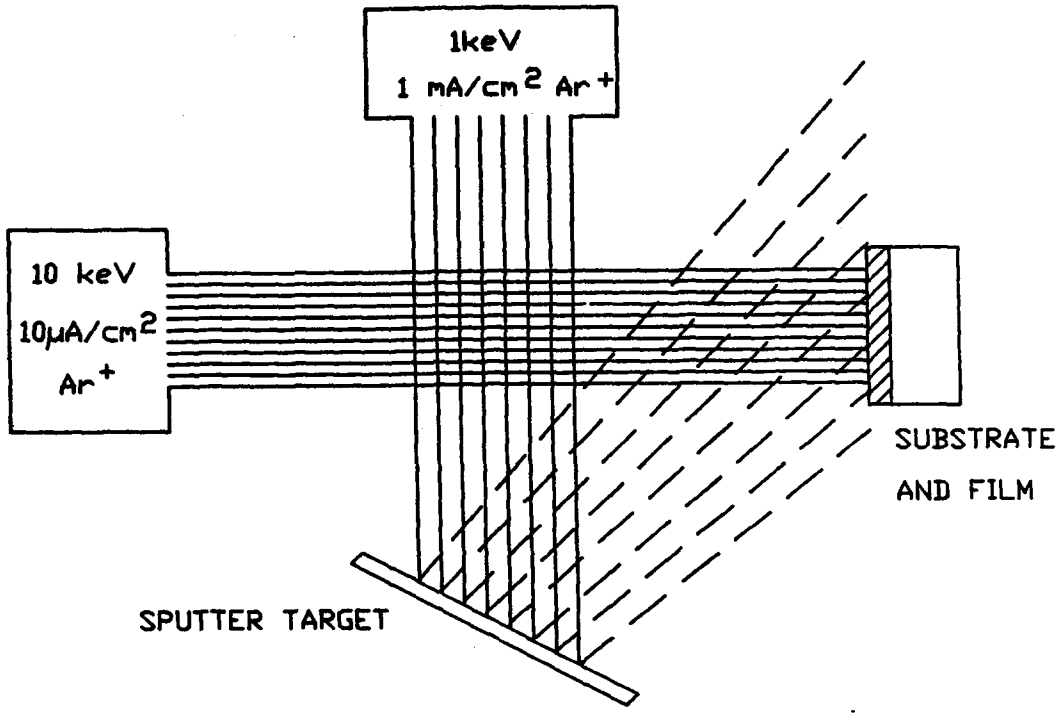


fig 1 Schematic of original DRM apparatus

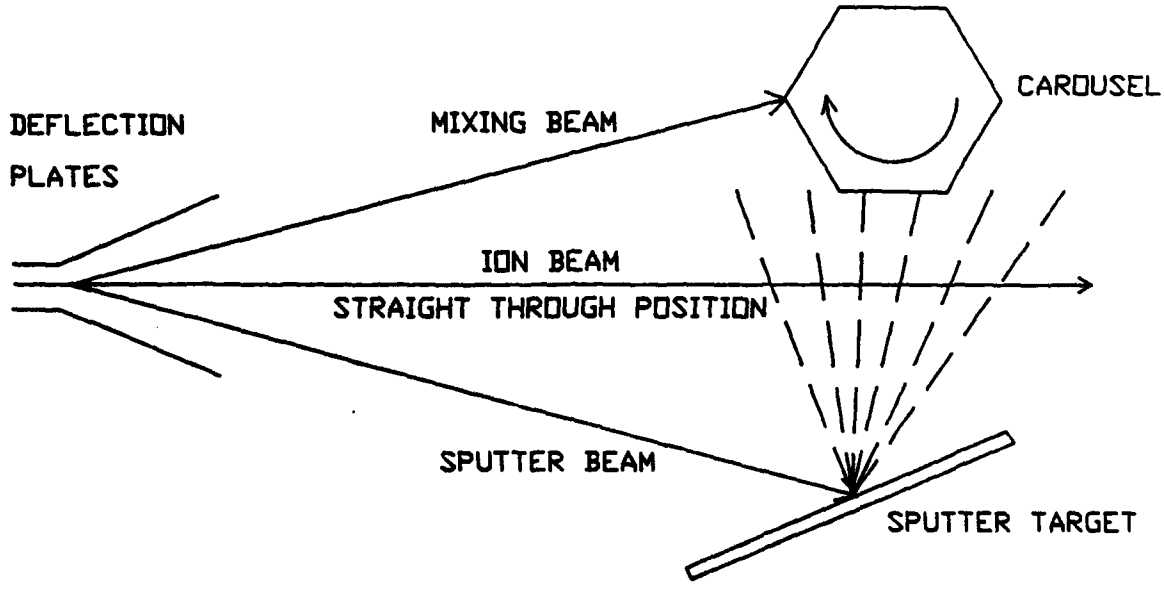


fig 2 Schematic of ion beam deflection system

of the recoil mixed samples was considerably better, with uniform $CoSi_2$ layers formed to a depth of 150 nm.

In the original experimental apparatus the film thickness was monitored using a quartz crystal, but due to the cross section of the beam it was not possible to continuously monitor the film thickness during DRM. The quartz crystal, but due to the cross-section of the beam it was not possible to continuously monitor the film thickness during DRM. The quartz crystal was therefore used to set up an initial dynamic balance on the assumption that it would remain constant during the period of processing. It was subsequently shown by RBS that this assumption was incorrect, and that the film thickness grew during the DRM process. This resulted in the movement of the position of maximum energy transfer away from the metal-silicon interface into the bulk film.

MODIFIED DYNAMIC RECOIL MIXING

A prototype system for the deposition of $CoSi_2$ for use in VLSI circuits has been developed and built based on a modified DRM technique as shown in figure 2. It makes use of constant film thickness monitoring allowing for a more accurate dynamic balance, a single ion source for both sputtering and mixing and a rotating carousel capable of holding five 3-inch silicon wafers.

It was decided to incorporate an electrostatic beam deflection system to cycle the beam between the sputtering and mixing modes because of its inherent simplicity compared with splitting the single beam into two subbeams. Using this new configuration, it is not possible to dynamically balance the new material arriving at the surface with the back-sputtered material. A new method was therefore devised where material is sputtered onto the substrate to a thickness slightly greater than the projected range of the mixing ions, the ion beam being deflected towards the sputter target. On reaching the pre-set thickness the beam is switched from the sputter to the mix mode where it remains until the sputtered film is reduced to a thickness slightly less than the projected range. The ion beam is then switched back to the sputter mode and the whole process repeated until the required dose is attained. Automatic control of the above process is achieved by the use of an optical film thickness instrument, which continually monitors the film thickness by measuring the light transmittance of the film deposited onto a glass slide.

The single ion beam DRM therefore relies on the thickness, t, of the film oscillating by an amount dt between two set points either side of the projected range. This has the overall effect of moving the position of maximum energy transfer, t, of the mixing ions, from the film, through the metal-silicon interface and into the substrate; i.e. from t - dt to t + dt (as with ion beam mixing). More metal is then sputtered onto the substrate and the process is repeated.

Sputter deposition rates for cobalt up to 4nm/minute have been achieved so that 150 nm of $CoSi_2$ should be formed in well under an hour. It is hoped that the prototype described here will lead to the development of a DRM machine capable of production wafer throughput.

OTHER SILICIDES

While the main application of this work has been to develop a DRM process for $CoSi_2$ formation; it is useful to make some comparisons with the formation of other silicides which have been produced using ion beam mixing. The main parameters comparing thermal and ion beam mixed growth appear in table 1 [6], [7], [8], [9] where T_c is the critical temperature below which no compound formation takes place. It can be seen that most silicide phases are formed at room temperature if ion beam mixing is used in contrast with thermally

THERMAL TREATMENT				ION BEAM MIXING			
Phase Formed	Tc (Thermal) °C	Activation Energy (Thermal) eV	Growth Vs Time t	Tc (Ion Beam) Induced °C	Activation Energy (Ion Beam) eV	Growth Vs Flux ϕ	Ion Species
$CrSi_2$	450	1.8	t	RT	0.18	ϕ	Cr
$MoSi_2$	540 - 1150	4.1	$t^{0.5}$	150	0.92	$\phi^{0.5}$	Ar
$TiSi_2$	600	1.5	$t^{0.5}$	RT	0.2	$\phi^{0.5}$	Xe
$TaSi_2$	650	3.7	t	RT	0.18	ϕ	Ar
$CoSi_2$	550			RT		$\phi^{0.5}$	Xe
CoSi	375 - 500	1.9	$t^{0.5}$	>RT			Xe
Co_2Si	350 - 500	1.5	$t^{0.5}$	RT	0.23	$\phi^{0.5}$	Xe
$NiSi_2$	>750			RT			Xe
NiSi	350 - 750	1.4	$t^{0.5}$	>RT	0.1	$\phi^{0.5}$	Ni,Xe
Ni_2Si	200 - 350	1.5	$t^{0.5}$	-100	0.1	$\phi^{0.5}$	Ni,Xe

activated compound formation. More recently, there has been evidence that the phase formed is dependent upon the nuclear energy loss at the interface [8]. The consequence of this result may be that optimisation of the mixing dose and any subsequent annealing cycle could be considered more complex than a simple ion beam mixed compound would imply. Further investigation is required to clarify these points.

ACKNOWLEDGEMENT

The work was funded by the Alvey Directorate as project VLSI 021.

REFERENCES

[1] Tsai, M.Y., Petersson, C.S., d'Heurle, F.M. and Maniscalco, V., Appl. Phys. Lett. 37 (1980) 295.

[2] Tsaur, B.Y., Chen, C.K., Anderson, Jr. C.H. and Kwong, D.L., J. Appl. Phys. 57 (1985) 1890.

[3] Fischer, G., Hill, A.E. and Colligon, J.S., Vacuum 28 (1978) 277.

[4] Colligon, J.S., Hill, A.E. and Kheyrandish, H., Vacuum 34 (1984) 843.

[5] Kozicki, M., PhD Thesis, University of Edinburgh (1984).

[6] Hamdi, A. and Nicolet, M.A., Thin Solid Films, 119 (1984) 357.

[7] Li, Wen-Zhi, Kheyrandish, H., Al-Tamimi, Z. and Grant, W.A., Nucl. Instrum. Methods, B19/20 (1987) 723.

[8] Schreter, U., So, F.C.T., Paine, B.M., Nicolet, M.A. and Johnson, W.L. in Hubber, G.K., Holland, O.W., Clayton, C.R. and White, C.W. (Eds), Ion Implantation and Ion Processing of Materials (North Holland, New York, 1984) pp 31.

[9] Kheyrandish, H., Colligon, J.S. and Stephens, G.A., Proceedings of 1986 Int. Vacuum Congress, Baltimore October 1986.

SIMULATION OF GATE-CONTROLLED DOUBLE-INJECTION SOI STRUCTURES APPLICATION TO MICROMAGNETODIODE SENSORS

G. DIMOPOULOS*, F. BALESTRA, A. CHOVET, M. BENACHIR AND J. BRINI

Laboratoire de Physique des Composants à Semiconducteurs (UA-CNRS 840)
Institut National Polytechnique de Grenoble
ENSERG, 23 Avenue des Martyrs, 38031 GRENOBLE Cedex, FRANCE

Among the sensors able to be integrated with VLSI technologies, the gated micromagnetodiode is one of the most attractive magnetic sensors due to its simple design, small size and good sensitivity.

Micromagnetodiodes have been fabricated on Silicon-On-Insulator technology. Experimental results concerning magnetic sensitivity, current-voltage characteristics and transconductance are compared to simple models which allow to understand more easily the physical behaviour of gated double-injection devices.

1. INTRODUCTION

It is known that an exact behaviour of gated devices made with silicon films on insulator could only be obtained by solving Poisson's equation in the structure. Numerical simulations of such solutions have recently been proposed [1], namely with the program ISIS-1 (one dimensional analysis along the direction : gate-SiO_2-Si film-Insulating substrate and back contact). This has led to a better understanding of electrical properties of enhancement-type and depletion-type thin film SOI MOSFET [2]. A very similar technology is used to realize magnetic microsensors, e.g. the gate-controlled micromagnetodiodes (insert of Fig. 1).

Magnetodiodes are double-injection p^+nn^+ structures which can be integrated with Silicon On Sapphire technology. The recently superimposed Si-poly gate appears to strongly influence, the device characteristics (Fig.1) [3].

In this paper, simple models are proposed in order to describe and analyze the properties (magnetic sensitivity, current-voltage characteristics, gate influence, ...) and the main parameters of the gated magnetodiodes.

2. THE MAGNETODIODE EFFECT AND RELEVANT PARAMETERS

Magnetodiode sensors belong to the family of semiconductor magnetic sensors based on carrier concentration under crossed electric and magnetic fields [4,5].

The carriers are injected into the n region (low-doped "base" of the device) by p^+(anode) and n^+(cathode) contacts. They form an e-h plasma which is deflected by a transverse magnetic field B (applied along z direction) towards one of the two asymmetric surfaces (in Silicon On Sapphire the carrier recombination velocity s_2 at the Si-Al_2O_3 interface is much higher than s_1 at the Si-SiO_2 interface; see insert of Fig.1). Therefore the carrier profiles in the Si film depend on the magnetic field polarity and intensity, which leads to variations of the mean resistance of the structure.

The current-voltage characteristics [6,7,8] are given by an extension of the well-known double-injection law, i.e. in the ohmic and semiconductor regimes :

$$J = q(n_o\mu_n + p_o\mu_p)\, V_D/L + (9/8)\, q\, \mu_n\mu_p\, |n_o - p_o|\, \tau_{eff}\, V_D^2/L^3 \quad (1)$$

*Alexander Onassis Foundation Scholar

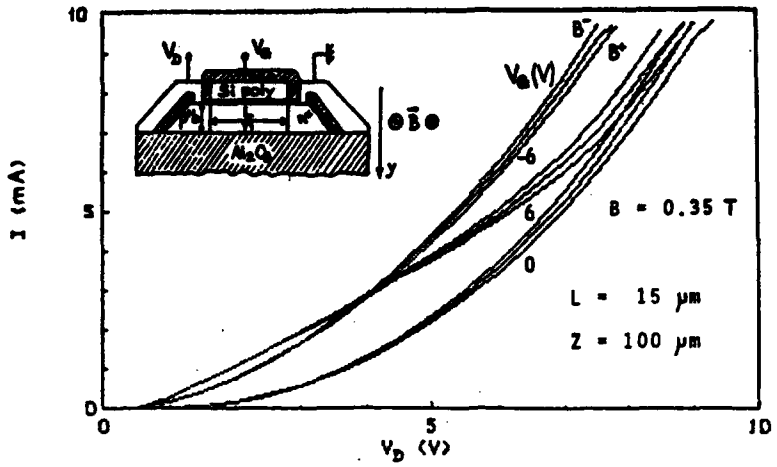


Figure 1 Configuration of Silicon-On-Sapphire p^+nn^+ gated micro-magnetodiode and typical experimental characteristics I *vs.* V_D

where

J is the mean current density across the device section,

V_D the voltage applied between anode and cathode,

L the length of the "base",

q the elementary electric charge,

n_o and p_o the equilibrium carrier concentrations,

μ_n and μ_p the carrier mobilities.

τ_{eff} which can be considered as an "effective" carrier lifetime is a key parameter, taking into account surface and magnetic field influences [6,7,8]. For low magnetic inductions ($\mu^2 B^2 \ll 1$), τ_{eff} is given by :

$$\tau_{eff} = \tau_v(1 - g_s + g_c E_x B) \qquad (2)$$

where τ_v is the carrier bulk lifetime.

g_s and g_c coefficients account respectively for the surface and the carrier concentration effects ; they depend on the recombination parameters (bulk and surface), the diffusion length and the device (film) thickness b.

3. A SIMPLE MODEL AND ITS LIMITATIONS

The following drastic simplification is used in the simplest model : the problem will be considered to be one-dimensional by taking in eq. (1) and (2) only the mean-values of parameters able to vary along y direction. Despite the fact that in such a case it is impossible to study the gate influence, it can be useful for analyzing the influence of technological and geometrical (n_o-p_o, μ, τ_v, b, L, device width Z) or electrical (V_D) parameters on quantities such as the current magnetic sensitivity S_I. We define, for a constant voltage V_D applied to the magnetodiode :

$S_I = \Delta I / \Delta B$, where $I = JbZ$ is the current across the device.

For two opposite directions of the magnetic field (+B and -B), we get from eq.(1) and (2) :

$$S_I = \frac{I_+ - I_-}{2B} =$$

$$= \frac{1}{2B}\,\frac{9}{8}\,q\,\mu_n\mu_p\,|n_o - p_o|\,(\tau_{eff+} - \tau_{eff-})\,\frac{V_D^2}{L^3}\,bZ$$

$$= \frac{9}{8}\,q\mu_n\mu_p\,|n_o - p_o|\,\tau_v\,g_c\,\frac{V_D}{L}\,\frac{V_D^2}{L^3}\,bZ \qquad (3)$$

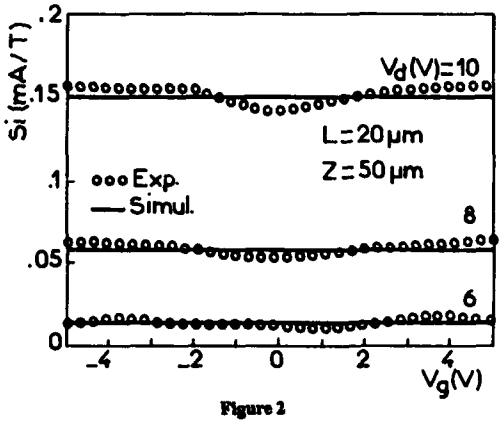


Figure 2

Experimental and simulated curves of current magnetic sensitivity S_I at different V_D for "long" micromagnetodiodes. Simulation parameters : $\mu_n = 200 cm^2/V.s$, $\mu_p = 90 cm^2/V.s$, $\tau_v = 20 ns$, $s_1 = 10 m/s$, $s_2 = 200 m/s$, Si film thickness $b = 0.6 \mu m$.

This calculation appears to be in correct agreement with experimental results for "long" S.O.S. magnetodiodes ($L \geq 15\mu m$) (Fig. 2) and the influences of the main parameters are well verified.

This simple model also shows that the sensitivity S_I would increase with the recombination velocity s_2 at Si-Insulator interface and would be inversely proportional to the recombination velocity s_1 at Si-SiO_2 interface.

In short magnetodiodes, the influence of the gate voltage V_G on S_I is stronger, and the assumptions corresponding to this simplified model are not always suitable so that there is not a good agreement between experiment and eq. (3) (Fig.3)

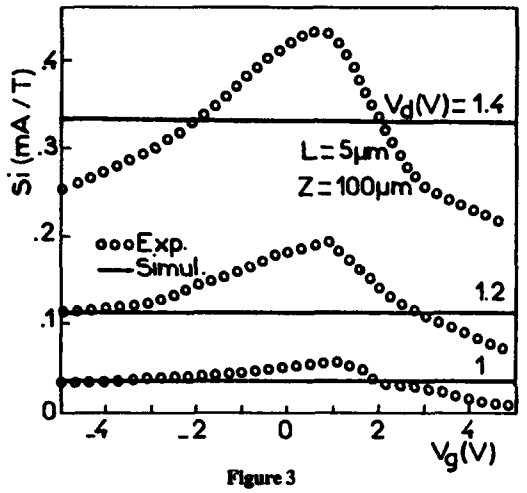


Figure 3

Experimental and simulated curves of current magnetic sensitivity S_I at different V_D for short micromagnetodiodes. Simulation parameters as in Fig.2

4. A MORE PRECISE MODEL

Here we consider that the carrier transport between anode and cathode as well as the magnetic field influence on $I(V_D)$ characteristics can be described by the nearly one-dimensional equations (1) and (2) but that the gate influence essentially consists in a modification of the equilibrium carrier densities $n_0(y)$ and $p_0(y)$ which govern the double-injection process.

Therefore the current across a gate-controlled double injection structure will be obtained by considering that the device tickness b is divided into m parts so that we have m elementary magnetodiodes connected in parallel :

$$I = Z \int J(y)\, dy = Z \Sigma J(y)\, b/m \qquad (4)$$

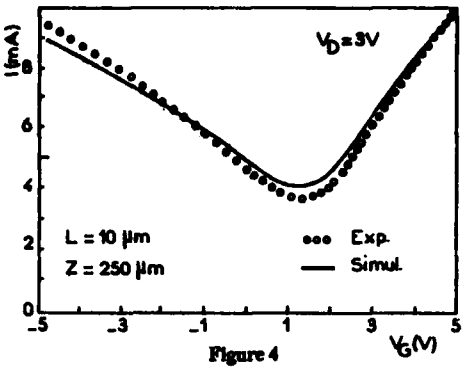


Figure 4

Experimental and simulated curves of current I vs. V_G for Silicon-On-Sapphire magnetodiodes.

Si film thickness $b = 0.6\mu m$. Simulation parameters : $\tau_{eff} = 0.8ns$.

Q_f/q : (SiO_2) : $< 10^{10}$ cm^{-2} ; (sapphire) : $7.10^{12} cm^{-2}$

N_{ss} : (SiO_2) : 5.10^{10} $cm^{-2}eV^{-1}$; (sapphire) : 10^{12} $cm^{-2}eV^{-1}$

μ_n linearly decreasing from 350 $cm^2/V.s$ (near SiO_2) to $70 cm^2/V.s$ (near sapphire); μ_p from 150 to 30 $cm^2/V.s$;

J(y) is calculated from $n_0(y)$ and $p_0(y)$, after solving the Poisson's equation by a numerical method : ISIS-1 program [1,2] uses a one-dimensional finite difference calculation with a nonlinear overrelaxation method and provides the potential $\phi(y)$ and the electron and hole densities $n_0(y)$, $p_0(y)$ inside the film for each gate potential V_G. Then the current characteristics $I(V_G)$ is deduced (for a constant voltage V_D).

Experimental and simulated curves for $I(V_G)$ and transconductance $G_g(V_G)$ are presented in Fig.4 and 5, and appear to be in fairly good agreement. It must be noticed that the simulated curves were translated by a constant voltage $V_{Go} \simeq +1V$ which exactly corresponds to the work-function difference ϕ_{MS} between a p^+-polysilicon electrode and n-Si film [9], not yet included in the program.

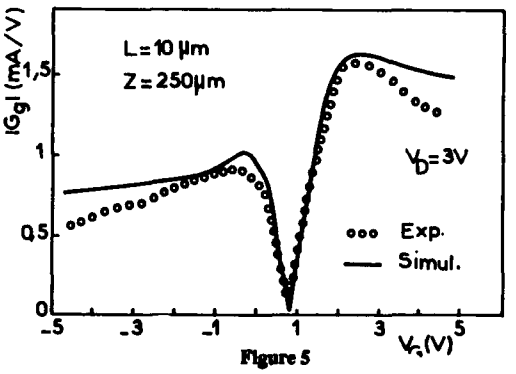


Figure 5

Experimental and simulated curves of transconductance G_g vs. V_G for Silicon-On-Sapphire magnetodiodes. Paramaters as in Fig.4

Also note that the parameters used for simulated curves correspond to typical values of physical parameters in S.O.S. devices and that the agreement between experiment and simulation is expected to be nearly perfect with a 2D numerical solution of the problem.

5. CONCLUSION

These simple models lead to the conclusion that in a gated magnetodiode, the double-injection phenomenon takes place as if there were a non-uniformly doped base with equilibrium carrier concentrations $n_0(y)$ and $p_0(y)$. Thus we get a better understanding of the dominant physical effects in the gate-controlled double-injection process. This will allow to optimize the technological parameters as well as the practical uses of such sensors.

ACKNOWLEDGEMENTS

Thanks are due to Dr. G. Ghibaudo for useful discussions and suggestions.

REFERENCES

[1] F. Balestra, J. Brini, P. Gentil, ESSDERC'84 (Lille 1984), in Physica 129B (1985) pp.296-300

[2] F. Balestra, J. Brini, P. Gentil, Sol. Stat. Electron. 28 (1985) pp.1031-1037

[3] G. Dimopoulos, DEA Report, INP - USM, Grenoble (1985)

[4] A. Chovet, S. Cristoloveanu, Rev. Phys. Appl. 19 (1984) pp.69-76

[5] H.P. Baltes and R.S. Popovic', Proc. IEEE, 74 (1986) pp.1107-1132

[6] H. Pfleiderer, Sol. St. Electron. 15 (1972) pp.335-353

[7] S. Cristoloveanu, Phys. Stat. Sol. (a) 64 (1981) pp.683-695, and Phys. Stat. Sol. (a) 65 (1981) pp.281-292

[8] S. Cristoloveanu, Doctorat ès Sciences, Thesis, INP Grenoble (1981)

[9] S.M. Sze, Physics of Semiconductors Devices, 2nd ed, John Wiley, New-York, 1981, p.397

OPTIMIZATION OF P IMPLANTED SILICON BOLOMETERS

E. Baciocco, C.Boragno, U.Valbusa

Dipartimento di Fisica, Via Dodecaneso 33, 16146 Genova, Italy

C.Bresolin, G.Pignatel[+]

SGS Microelettronica, Agrate Brianza, Milano, Italy

We have measured electrical resistance R as function of temperature T in the range 2 ÷ 20 K of several P implanted Silicon bolometers with the aim of optimize the characteristics of the bolometers. We have investigated the R(T) behavior of several samples obtained by varying implant conditions, depth profile and annealing.

1. INTRODUCTION

P implanted Silicon bolometers have been used recently for detection of molecular beams, infrared radiation (1), alpha particles (1),(2),(3), X rays (4) and ballistic phonons (5) , showing a large use in different fields of physics and chemistry such as astrophysics, surface science, atomic and molecular physics, laser spetroscopy, nuclear and subnuclear physics.

Silicon P doped low temperature bolometers, doped with the method of the ion implantation, have been realized by several authors (1), (6). This method offers several advantages over the diffusion-doping since this technique allows to create thin doped films which in consequence have very fast responce time (5).

The aim of the present work is to gain insight in the physics of the variable range hopping conduction on which is based the working principle of the bolometer and at the mean time to investigate the influence of the thermal treatment of the sample on the electrical conduction.

2. EXPERIMENT

2.1. Experimental set up

The apparatus to measure the electrical resistance R of Silicon P doped samples in the range 1-20 K it will be described elsewhere (7). Here we limit ourself to give a short description of the method used to measure R(T). It is based on the measurement of the sample resistance by using a standard lock-in amplifier. This allows to measure resistance in the range 10 M ohm -1 K ohm dissipating 1 nWatt- 0.1 pWatt in the sample in order to avoid self heating. The sample is inserted in a chamber in poor thermal contact with a liquid helium bath. The temperature of sample varies at a rate of 10^{-3} K sec^{-1} and it is measured by a carbon resistance.

The lock-in amplifier is interfaced with a

[+]Present address: Dipartimento di Ingegneria Elettronica, Università di Bari, Bari, Italy

PC IBM AT which records R(T). The method allows to measure 8 samples at the same time with an accuracy of 1% in the temperature reading and of 0,1% - 1% in the resistance depending on the value of the resistance itself.

2.2. Sample preparation

Single crystals of Si were prepared with P donor concentrations ranging from 3.7 - 2.0 10^{18} cm^{-3} near the critical concentration of the metal-insulator transition N_c=3.74 10^{18} cm^{-3} (8). The samples were obtained by implanting different doses of P at different energies in order to obtain a box profile.

A typical profile is reported in Figure 1. It has been computer simulated by using the program SUPREM III.

The resulting box profile gives an average concentration of n=2.0 10^{18} cm^{-3}.

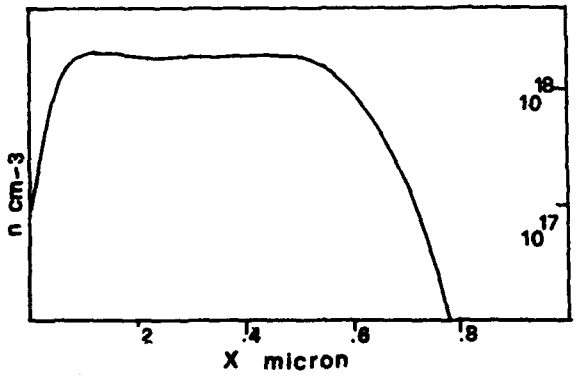


FIGURE 1

Depth from surface X for n=2.0 10^{18} cm^{-3}.

Similar calculations have been carried out for all the sample reported in this paper.

All the samples are then activated by annealing at 920 °C in N_2 for 15 min and then in O_2 at the same temperature for 15 min.

3. MEASUREMENTS

Figure 2 illustrates observed resistance R as a function of $T^{-1/4}$ in the region 2-20 K for some samples. The curves follow the temperature dependance derived by Mott (9)

$$R = Ro \exp (To/T)^{1/4} \qquad (1)$$

Best fit procedure on the data allowed to determine experimental values of To and Ro.

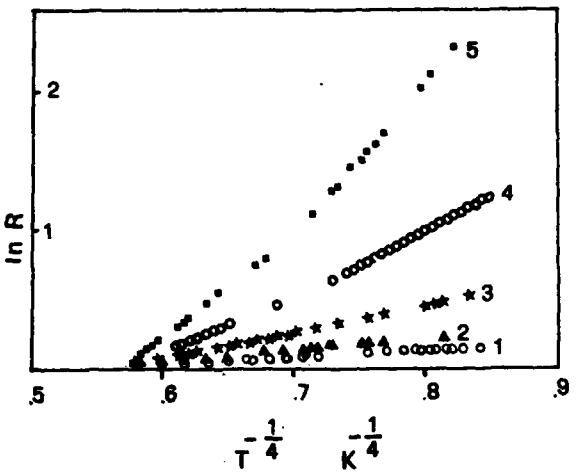


FIGURE 2

Plot of ln R versus $T^{-1/4}$ according to the Mott law. All curves have been normalized at R(10 K).

Figure 3 reports R(T) measurements for samples of same concentration n=2.2 10^{18} cm^{-3} which have received different thermal treatments. Sample 1 has been annealed at 600 °C for 1

hour in N_2 and subsequently at 920 °C for 15' in N_2 and 920 °C for 15' in O_2.

Sample 2 has been annealed only at 920 °C for 15' in N_2 and 920 °C for 15' in O_2 and sample 3 at 600 °C in N_2. By looking in Figure 3 it appears that sample 2 shows a temperature behavior with a slope lightly steeper than sample 1 and both result much steeper than sample 3 showing clearly the importance of annealing and leading to the conclusion that the standard annealing procedure described in section 2 is very satisfactory.

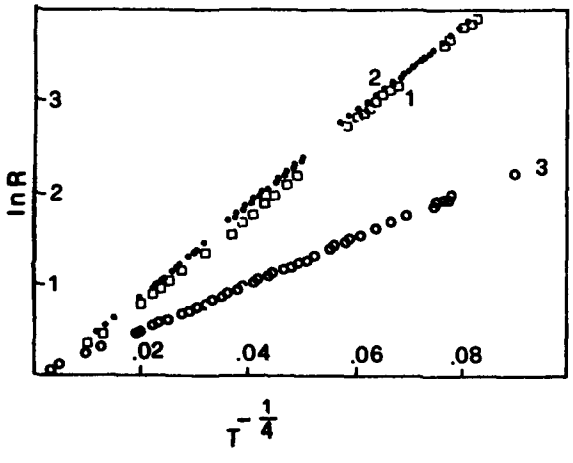


FIGURE 3

Plot of ln R versus $T^{-1/4}$ for the same sample n=2.2 10^{18} cm^{-3} and different annealings. All curves have been normalized at R(20 K) and T=20 K.

4. SCALING LAW

It is well known that the responsivity of a bolometer depends by the temperature coefficient of resitsance $a = (1/R)\frac{dR}{dT}$

In our case a can be expressed in term of To by the Mott law resulting

$$a = -\frac{To^{1/4}}{4} T^{-5/4}$$

The To values obtained by the data of Figure 2 and others not reported can be scaled with the concentration ratio $\frac{n}{n_c}$ as reported by (10) following

$$To = \frac{b}{k\, N(E_F) l_o^3} (1 - \frac{n}{n_c})^{3P} \qquad (2)$$

where b is a numerical value, $N(E_F)$ is density of states at the Fermi energy, k the Boltzmann constant, n_c the crytical density, l_o the effective Bohr radius of the donor and P a parameter which according to the scaling theory of localization (11) is P = 1.

Equation 2 can be taken as a scaling law for To. Figure 4 reports the values of ln To versus $\ln(1 - \frac{n}{n_c})$.

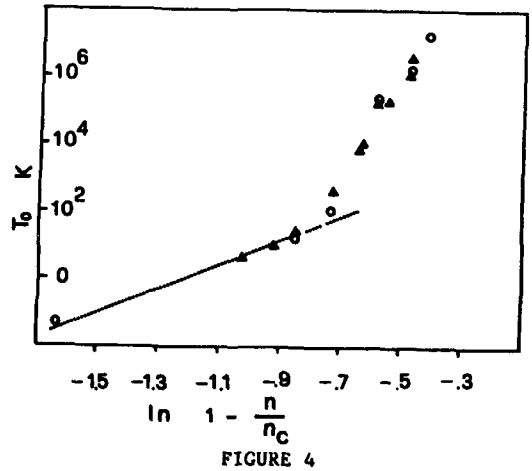


FIGURE 4

ln To versus ln $(1-n/n_c)$.

Near the critical density, the curve follows the behavior of equation (2) with an experimental value of $3P = 3.3 \pm 0.05$. Similar scaling law has been found in Si:As near the

critical concentration (10).

ACKNOWLEDGEMENTS

This work has been supported by GNSM of CNR and by CISM of MPI of Italy.

We would like to thank M.Sassetti for useful discussion and A.Gussoni for technical assistantship.

REFERENCES

(1) Boragno, C., Valbusa, U., Gallinaro, G., Bassi, D., Iannotta, S. and Mori, F., Cryogenics, 24 (1984) 681.

(2) Niinikoski T.O., Rijllart A., Alessandrello, A., Fiorini, E. and Giuliani A., Europhys. Lett. 1 (1986) 499.

(3) Coron, N., Dambier, G., Focker, G.J., Hansen, P.G., Jegoudez, G., Jo, B., Leblanc, J., Moalic, J.P., Ravn H.L., Stroke, H.H. and Testard, O., Nature 75 (1985) 314.

(4) McCammon, D., Moseley, S.H., Mather, J.C. and Mushotzky, R.F., J. Appl. Phys., 58 (1984) 1263.

(5) Boragno, C., Valbusa, U. and Pignatel G., Appl. Phys. Lett. 50 (1987) 583.

(6) Downey, P.M., Jeffries, A.D., Meyer, S.S., Weiss, R., Bachner, F.J., Donnelly, J.P., Lindley, W.T., Mountain, R.W. and Silversmith, D.J., Appl. Opt. 23 (1984) 910.

(7) Baciocco, E., Boragno, C., Valbusa, U., Bresolin, C. and Pignatel, G., to be published.

(8) Rosenbaum, T.F., Milligan, R.F., Paalanen, M.A., Thomas, G.A., Bhatt, R.N. and lin, W. Phys. Rev. B 27 (1983) 7509.

(9) Mott, N.F., Phyl. Mag. 19 (1969) 835

(10) Shafaram, W.N. and Castner, T.G., phys. Rev. B 33 (1986) 3570.

(11) Abrahams, E., Anderson, P.W., Licciardello, D.C. and Ramakrishnan, T.V., Phys. Rev. Lett. 42 (1979) 673.

LARGE AREA SILICON DETECTORS FOR CALORIMETRY IN HIGH-ENERGY PHYSICS

B. Passerini, M. Zambelli, P.E. Zani

ANSALDO - Unita' Semiconduttori
Via N. Lorenzi 8, 16152 Genova, Italy

This paper presents the preliminary results obtained in the development of large area p-i-n diodes to be used as particles and radiation detectors in high-energy physics experiments. Large area room-temperature operated silicon detectors widen the potential field of application for solid state detectors in high-energy physics. Large equipments, like electromagnetic and hadronic calorimeters, could be built, in the near future, using large active areas of silicon to the place of more traditional detectors, like scintillators, gas chambers, cryogenic and room temperature liquids.

1. INTRODUCTION

Silicon detectors are a major instrument in high-energy physics experiments [1]. They are very attractive because they can achieve high spatial resolution, have a fast response time, can operate at room temperature, in strong magnetic fields, in vacuum, and are adequate for arrangements with particular geometric constraints. In the past silicon counters have been employed mainly as vertex detectors, but now they are expected to find wide application in calorimetry. This application requires a large amount of active silicon area in order to cover the whole solid angle around the point of collision of the incident particles. In order to arrange tens or hundreds of square meters of silicon area, any single detector should achieve the maximum possible area and several detectors should be arranged in a module.

A silicon radiation detector is basically a planar p-i-n diode operated in reverse bias. The incident radiation, gamma-rays, x-rays or charged particles, produces electron-hole pairs which are swept to the n^+ and p^+ contacts by the applied electric field. Such a p-i-n diode, to be suitable as a high quality radiation detector, must have a sensitive region free of charge trapping centers, contacts thin enough to allow the penetration of the incident radiation, low bulk and surface defects to limit at very low levels the reverse bias leakage current. The detectors presented in this work are actually the largest silicon junction detectors up to now developed for radiation and charged particles detection. These devices have been used at the end of 1986 by the SICAPO collaboration # to build a prototype of an electromagnetic calorimeter [2].

Next section presents the fabrication process and some aspects of the physics of operation. Afterwards the application in calorimetry is discussed.

Universities of Hamburg, McGill, TelAviv, and Italian Institute for Nuclear Physics INFN (Genoa, Florence, Milan, Trieste).

2. FABRICATION TECHNOLOGY AND OPERATION

The starting material for detectors fabrication is high resistivity single crystal silicon. It is usually available from silicon suppliers as a by-product of the material used by the power devices industry. Common specifications are float-zoned crystals, <111> oriented, n-type (neutron transmutation doped, in some cases). Up to now however, owing to a marginal interest of silicon suppliers in the detectors business, no deep understanding of the crystal growth parameters affecting the detector performances exists, and a "detector grade silicon" cannot be specified.

In this work silicon wafers 100 mm in diameter and 400 μm thick have been used. The first step in the fabrication process is the growth of a layer of silicon dioxide 1 μm thick, performed in a standard furnace in steam and chlorine atmosphere at a temperature in the range 950-1050 C. This oxide layer is used as a mask for the subsequent ion implantation of the p+ anode region. To this purpose windows are opened in the oxide to define the geometry of the detector with a standard photolitographic process. Boron ion implantation on the detector front and phosphorus implantation on the back are carried out to obtain the p^+-n-n^+ structure. In both cases low energy implants (20-40 keV) are used with a total dose in the range $1x10^{14}$-$1x10^{15}$ cm^{-2}. An annealing process at 600 C is performed in order to reconstruct the crystal lattice removing the damage produced by the implantation process. Aluminum is then evaporated on the two sides of the silicon wafer and patterned on the front, according to the detector geometry, with a second photolitographic process. A last optional step, the passivation coating, is performed only when full environmental protection of the detector is required. The passivation coating is a layer of polyimide resin, which covers both the silicon dioxide and, apart the contact region, the metallized surface of the detector front. Fig. 1 shows a picture of a finished detector. In this work a geometry with two detectors with a trapezoidal shape and 27 cm^2 of implantation area have been realized on wafers 100 mm in diameter and 400 μm thick.

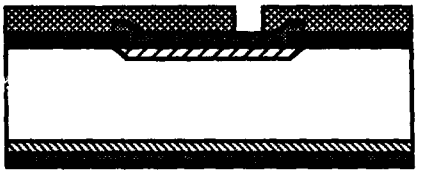


FIGURE 1

Cross sectional view of a finished detector.

The active area is extended laterally beyond the implanted area owing to the junction curvature at the edge. This extended active area, A_E, is to a first approximation, proportional to the depleted region X_d [cm]:

$$A_E = \lambda X_d P \ [cm^2]$$

where P is the juntion perimeter in cm and λ is the ratio between the lateral extension of the active area and the depleted layer width. The detector capacitance is

$$C_D = 1.0359[(A_j+A_E)/X_d] \ [pF]$$

where A_j is the junction area in cm^2. For n-type bulk silicon the extension of the depleted layer is

$$X_d = 0.529x10^{-4} \sqrt{[\rho(V+V_B)]} \ [cm]$$

where ρ is the detector resistivity in ohm.cm, V_B is the built-in voltage, and V is the reverse bias voltage. From the preceeding equations one obtains

$$C_D = 1.958 \times 10^4 (A_j/\sqrt{[\rho(V+V_B)]}) + 1.0359\ P\lambda \quad [pF]$$

where $1.0359\ P\lambda$ is the correction to the standard formula of C_D. This term, is found to be small (less than 1%) [2]. Fig. 2 shows the leakage current as a function of the externally applied voltage for different detectors of the same process batch. The implantation doses and energies were in this case $1 \times 10^{14} cm^{-2}$ and 40 keV both for boron and phosphorus. Fig. 3 shows the capacitance vs. the applied voltage in the form $1/C^2$. Again curves of different detectors in the same process batch are reported. If the correction factor in the relation defining C_D is neglected, for non full depletion conditions the $1/C^2$ vs. V relationship is given by

$$1/C^2 = \rho(V+V_B)[1/(1.958 \times 10^{14} A_j)]$$

The slope of the curve is then proportional to the resistivity of the bulk material. The point beyond which saturation occurs gives the voltage at which full depletion is achieved. As can be seen a relatively large variation of the bulk resistivity of silicon is found. This brings about a corresponding variation in the voltage to be applied to reach full depletion. Table I gives the full depletion voltage values and the corresponding leakage current, together with the computed values of the silicon bulk resistivity of the detectors considered here. Relatively low resistivity material is expected to have better performances than high resistivity material above all with regard to radiation damage effects [3].

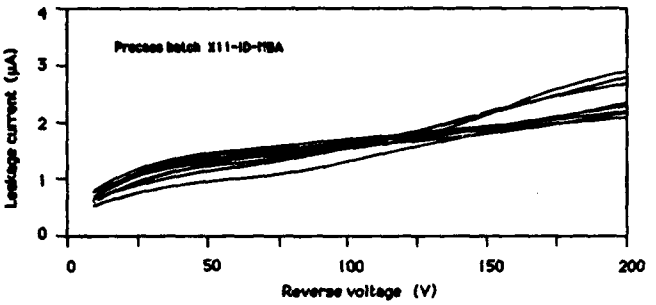


FIGURE 2
Leakage current in reverse biased detectors. The curves refer to different devices in the same process batch.

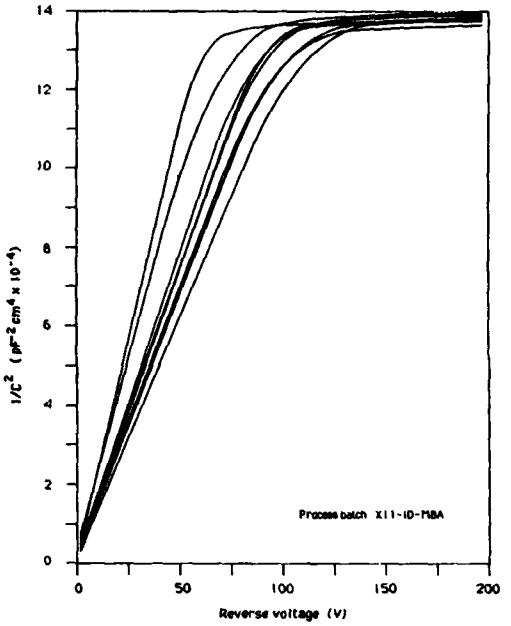


FIGURE 3
Capacitance-voltage relationship for the same detectors of Fig. 2.

Sample no.	Full depl. voltage (V)	Leakage current at FD (μA)	Si bulk resistivity (Ω cm)
1	122	1.8	4.68
3	100	1.6	5.69
8	76	1.3	7.52
12	101	1.7	5.65
16	111	1.7	5.17
17	95	1.5	5.99
18	111	1.7	5.17
23	71	1.1	8.00

TABLE I
Essential characteristics of the detectors shown in the preceeding figures.

3. APPLICATIONS TO CALORIMETRY

In calorimetric measurements performed in high-energy physics the energy of the incident particles must be evaluated. To this purpose a silicon calorimeter consists essentially in a sandwich of active sampling planes (silicon p-i-n diodes) and layers of a high atomic number passive absorber (usually uranium or tungsten). According to the kind of particles to be detected, calorimeters are referred to as electromagnetic and hadronic. The first ones are specially designed to detect leptons and electromagnetic radiation, while the second ones are dedicated to hadron identification. The purpose of the development of the detectors presented in this work is to build a full hadronic calorimeter as a test equipment to evaluate the feasibility of large scale applications. Currently a prototype of a Si/U calorimeter is in an advanced stage of development. It will consist of 20 silicon mosaic planes with an active area of about 500 cm^2 per mosaic containing 18 trapezoidal detectors each. These planes are interspaced with layers of uranium 5 mm thick. It is planned by the SICAPO collaboration to develop a full Si/U calorimeter, consisting of 130 silicon sampling planes (for a total active area of about 6.5 m^2). A silicon sampling plane is made of 18 trapezoidal detectors [4]. These modules are built using two fiberglass sheets in order to sustain mechanically the detectors and to bring the electrical connections to both their
junction and rear sides. The detectors are attached to the fiberglass sheets by silicon rubber and electrically connected to the metal paths by conductive epoxy resin. The contacts are brought to an external fan-out, which can be either a standard connector or a flexible capton cable.

4. CONCLUSIONS

Large-area p-i-n diodes fabricated on high resitivity silicon wafers have been used as particle detectors in a prototype of an electromagnetic calorimeter tested at CERN. In order to enable the use of silicon detectors for hadron calorimeters, mosaic modules consisting of 18 trapezoidal detectors, with an active area of 27 cm^2 each, were developed, assembled and tested. In the performed investigations no physical deterioration was observed in the detector characteristics. The results obtained confirm that silicon sampling calorimeters can fullfill the requirements of linearity, granularity, compactness, long-time stability and reliability needed in colliding beam machines experiments [2,5].

ACKNOWLEDGEMENTS

The authors are greatly indebted to Dr. P.G. Rancoita of INFN for all the experimental work done on measurement and application of detectors.

REFERENCES

[1] Rancoita, P.G., Journal of Physics (1984) 299.

[2] Rancoita, P.G., and Seidman, A., CERN Report EP/86-113 (1986)

[3] Borgeaud, P., McEwen, J.G., Rancoita, P.G., and Seidman, A., Nucl.Instr. and Meth. A (1983) 211.

[4] Pensotti, S., Rancoita, P.G., Seidman, A., Vismara, L., and Zambelli, M., CERN Report EP/86-103 (1986).

[5] Barbiellini, G., Cecchet, G., Hemery, J.Y., Lemeilleur, F., Leroy, C., Levman, G., Rancoita, P.G., and Seidman, A., Nucl. Instr. and Meth. A (1985) 316.

A MAGNETIC FIELD SENSOR USING A GRADED GATE POTENTIAL

B. S. Gill and E.L. Heasell.

Department of Electrical Engineering,
University of Waterloo,
Waterloo, Ontario,
Canada. N2L 3G1

Conventional, split-drain MAGFETs have a relatively low sensitivity. The low sensitivity can be attributed to properties inherent to the traditional FET structure. A novel device structure, designed to overcome these shortcomings, is presented. In this device the gate voltage varies linearly along the gate. Measured sensitivities for both dual-drain and triple-drain devices are reported. The performance of the device is superior to that of the conventional MAGFET devices.

1. INTRODUCTION

A variety of semiconductor, magnetic-field sensors have been described in the literature. Baltes [1] has given a recent, comprehensive review of such devices.

Sensors fabricated using standard, silicon processing are attractive since they would permit the incorporation of ancillary circuitry on the chip.

MOSFET devices appear to offer the advantage of thin, conducting channel regions, in which a larger Hall effect might be anticipated. However, the Lorentz force is a function of carrier velocity rather than the actual channel thickness. In the split-drain MAGFET [2], [3] carrier redistribution can occur only in the immediate vicinity of the drains. In the channel a uniform Hall field requires zero space-charge. In the drain region the effect of velocity saturation reduces the Hall angle and degrades the redistribution of channel current.

In an attempt to overcome some of these disadvantages we have fabricated MOSFET-like structures having a resistive gate (undoped polysilicon), with ohmic end contacts. A bias is applied across the gate, creating a uniform, *aiding field* along the channel. Simple device theory shows that the electric field along the channel can be made almost constant, at a value set by the differential gate bias.

Carrier collection is performed by long, lateral drain contacts, disposed symmetrically at the sides of the channel. In some devices a third drain is added to collect *un-deflected* carriers.

In the two-drain devices, the lateral drain diffusions are reverse biased to a potential greater than that at any point along the channel. Charge accumulation, at the edge of the channel, is prevented, no Hall field is established, and carriers are collected over a significant fraction of the channel.

We present results on both dual-drain and triple-drain Graded-gate (GFET) structures. Sensitivities observed for these devices are significantly better, than those for corresponding MAGFET devices.

2. DEVICE GEOMETRY

A layout for a typical GFET structure is shown in Figure 1. The drain contacts are located along the channel; this increases the area over which carrier collection can take place and prevents the development of the Hall field (i.e., the device is based purely on carrier deflection).

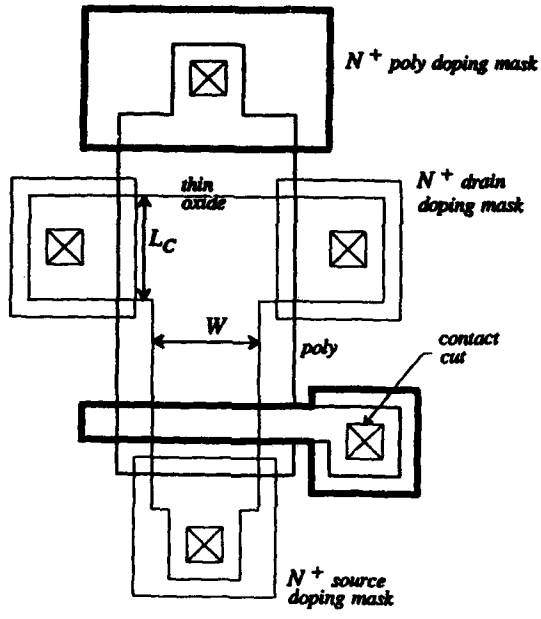


Figure 1. *Typical layout of the Graded-gate FET.*

There is a diffused gate-contact at each end of the device, with undoped polysilicon forming most of the gate. By biasing the gate contacts separately the resulting voltage gradient along the gate produces a more uniform longitudinal electric field along the entire length of the channel.

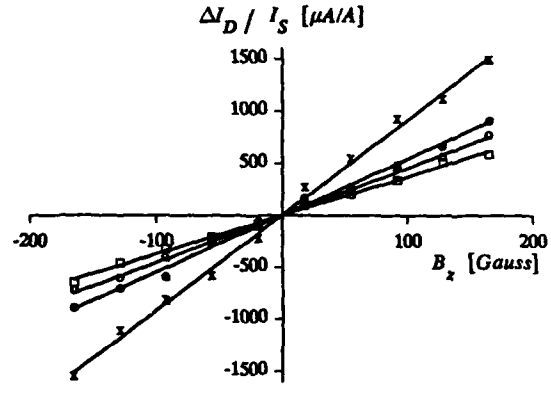
The width of the device W is taken as the minimum width of the channel, and the length L is defined to be the minimum length the carriers must traverse to be collected by one of the lateral drains. We introduce another parameter for this device L_C, the length in the channel direction over which the collection of the carriers can take place.

3. DEVICE FABRICATION

Fabrication of the devices was carried out on <100> p-type substrate with a resistivity of $1\Omega \cdot cm$. As we required the polysilicon over most of the channel to remain undoped, a standard NMOS process could not be used. The n^+ source and drain diffusions were performed prior to the deposition of the gate-polysilicon. An extra mask level was required to mask the majority of the poly, while the ends of the poly were doped to provide ohmic end-contacts. The devices had a threshold voltage of $0.6V$ and an oxide thickness of $825\mathring{A}$.

4. DEVICE PERFORMANCE

The application of a magnetic field perpendicular to the current flow produces an imbalance in the two drain currents. This assymetry in the drain currents provides a measure of the strength of the magnetic field. In Figure 2, we plot the relative current imbalance $\Delta I_D / I_S$ as a function of the magnetic field for various values of the differential gate-bias ΔV_g. The geometry of this GFET is $W = 20\mu m$, $L = 40\mu m$, and $L_C = 60\mu m$. The current imbalance is linear in B_z, allowing for experimental error and the influence of noise.



LEGEND		
SYMBOL	$\Delta V_g = V_{g_L} - V_{g_o}$ [V]	*SENSITIVITY* [mA /AT]
□	0	37.3
o	5	45.5
●	10	54.5
x	15	91.5

Figure 2. *Measured response of the GFET to an applied magnetic field with a positive voltage gradient (i.e., $\Delta V_g \geq 0$) along the gate (V_{g_L}=20V and V_{D1}=V_{D2}=20V).*

In all sensors there exists an offset signal, i.e., a current imbalance in the absence of a magnetic field. This can be attributed to manufacturing defects. To facillitate comparison, this offset signal has been removed in Figure 2.

The relative sensitivity of the devices is defined as:

$$S_r = \frac{1}{B_z}\frac{I_{D2} - I_{D1}}{I_{D2} + I_{D1}} .$$

I_{D1} and I_{D2} are the currents in each of the lateral drains. From the table in Figure 2 we see that the sensitivity of the GFET increases with increasing gate-voltage gradient. In Figure 3 we plot the sensitivity of the GFET as a function of the differential bias on the gate: $\Delta V_g = V_{g_L} - V_{g_o}$. V_{g_L} and V_{g_o} respectively denote the gate voltage at the drain end and the source end of the channel. Both positive (i.e., $\Delta V_g > 0$) and negative (i.e., $\Delta V_g < 0$) gate-voltage gradients were considered. For $\Delta V_g \geq 0$, V_{g_L} was fixed at $20V$ and V_{g_o} was varied from $20V$ down to $1V$. Similarly, for $\Delta V_g \leq 0$, V_{g_o} was fixed at $20V$ and V_{g_L} was varied from $19V$ down to $-1V$.

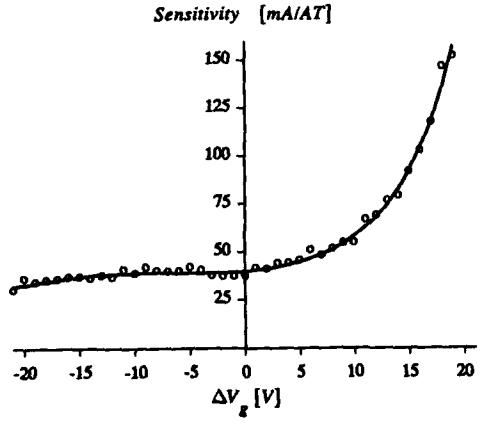
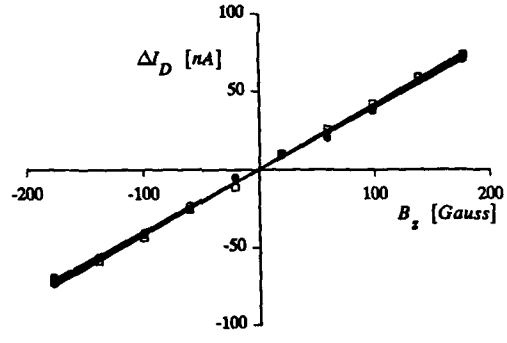


Figure 3. Detailed plot of the relative sensitivity of the GFET as a function of the gate-voltage gradient, $\Delta V_g = V_{g_L} - V_{g_o}$.

From Figure 3, we observe that the sensitivity improves considerably as we increase ΔV_g (for $\Delta V_g > 0$) and remains fairly constant for $\Delta V_g < 0$. The maximum measured sensitivity was five times larger than that obtainable with $\Delta V_g = 0$.

Triple drain GFET structures were also considered. The aim in these devices was to collect the undeflected carriers in the central drain. The measured response of a triple drain structure operated with $\Delta V_g = 0$ is shown in Figure 4. The geometry of this device is similar to the dual-drain device of Figure 2; the distance from the source to the center drain was $140\mu m$. The bias on the lateral drains is intermediate, between the source and centre-drain potential. The lateral drain diffusions act as collectors, close to the source end of the channel and as sources, close to the drain end. Extremely high sensitivities are observed in the



LEGEND			
SYMBOL	V_{D1} [V]	V_{D2} [V]	*SENSITIVITY* [mA /AT]
●	0.998	0.991	2,330
□	0.940	0.933	3,480
o	0.961	0.953	66,100

Figure 4. Experimental data for the triple-drain GFET with $\Delta V_g = 0$ ($V_{g_L} = V_{g_o} = 5V$), $V_{D1} \approx V_{D2} \approx 1V$, and $V_{Dmiddle} = 5V$. The extremeley high sensitivities are achieved as a result of the central drain collecting the undeflected current.

three-drain devices. However, this is achieved at the expense of great difficulty in establishing balanced current between the lateral drains. This makes the triple-drain GFET structures impractical in the present form.

5. CONCLUSIONS

In the conventional split-drain MAGFETs, the Hall angle is a decreasing function of the longitudinal electric field. The devices are usually operated beyond pinch-off where the longitudinal field is large. Hence, devices which rely soley on the deflection of carriers in the vicinity of the drain are not fully optimized.

A new device, the Graded-gate FET (GFET) designed to overcome these problems, was presented. The differential gate bias allows control of the surface potential and hence the longitudinal electric field. Also, carriers are collected in the lateral drain region to prevent the establishment of the Hall field. The sensitivity of these devices increased with increasing postive gate-voltage gradient. Negative gate voltage gradients neither significantly enhanced or degraded the magnetic sensitivity.

REFERENCES

[1] H.P. Baltes and R.S. Popović, *Integrated Semiconductor Magnetic Field Sensors*, Proc. IEEE, 74, No. 8, 1107-1132, Aug. 1986.

[2] P.W. Fry and S.J. Hoey, *A Silicon Magnetic Field Transducer of High Sensitivity*, IEEE Trans. Electron Devices, **ED-16**, No. 1, 35-39, Jan. 1969.

[3] D. Misra, T.R. Viswanathan and E.L. Heasell, *A Novel High Gain MOS Magnetic Field Sensor*, Sensors and Actuators, **9**, 213-221, Sept. 1986.

NUMERICAL MODELLING OF MAGNETIC FIELD SENSITIVE MOSFET

Ho Yie Wei Tongli & Shen Kechang

Microelectronics Center
Nanjing Institute of Technology, China

Abstract-A two dimensional numerical modelling of Magnetic-Field-Sensitive MOSFET (MAGFET) is presented and has a small computation cost. The whole operation region of MAGFET is divided into two parts (normal and subthreshold region) and numerically simulated respectively. A conclusion which MAGFET has highest sensitivity when W/L=0.8 is obtained and is verified by experiment.

1. INTRODUCTION

Along with the rapid progress of microelectronics industry, semiconductor sensor plays more and more important role in many fields. As an entrance transducer that converts the magnetic field into an electric signal, the Magnetic-Field-Sensitive MOSFET (MAGFET) offers advantages of high sensitivity, lower power dissipation and compatibility with an advanced standard MOS or CMOS IC technology. In view of that MAGFET is governed by the complexity of the interaction of semiconductor bulk effect with boundary conditions and magnetic field. A two dimensional numerical modelling of MAGFET is required to analyze the device multidimensional effect.

As it is known, the general approach of solving the nonlinear system of partial differential equation in semiconductor is coupled or decoupled solution method. In comparision with coupled approach, the decoupled approach requires much less computer memory. In view of that, the decoupled solution method is used. Furthermore, in the numerical modelling of MAGFET, the CPU time required by decoupled approach nearly is an exponential function of bias voltage when gate bias exceeds the threshold voltage. To avoid the trouble, the nonlinear partial differential equation system is simplified into a Laplace's equation in normal region and is numerically solved. In subthreshold region, the decoupled solution method is still employed.

2. MODEL AND NUMERICAL ALGORITHM

2.1. Subthreshold Region

The structure of n-channel MAGFET is shown in Fig.1. It has a split-drain structure. The z-axis is assumed to be parallel to the magnetic induction $\vec{B}$ (see Fig.2), viz., the magnetic field is perpendicular to the device surface.

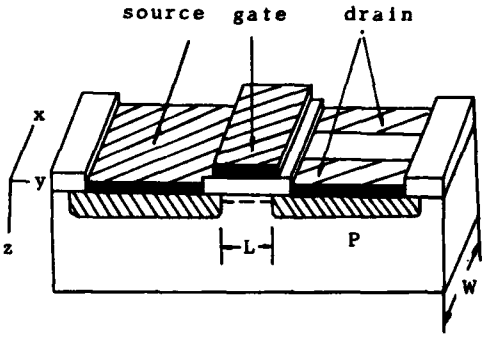


Fig.1 Structure of MAGFET

Then the distribution of carrier and electric potential in the device depend only on x and y.

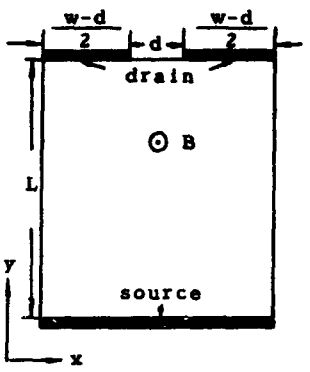


Fig.2 Geometry of the Simulated Device

The model equation in semiconductor is

$$\frac{1}{q}\nabla \cdot j_p + R = 0 \qquad (1)$$

$$\frac{1}{q}\nabla \cdot j_n - R = 0 \qquad (2)$$

$$\Delta\psi = \frac{q}{\varepsilon_{si}}(p - n + N) \qquad (3)$$

where R is the bulk recombination rate, N is the ionized net doping density, n and p are the electron and hole concentration respectively. In a stationary and not too large magnetic field, the current density equation reads as follows

$$\vec{j}_n = q(u_n n\vec{E} + D_n\nabla n) - u_n^*(\vec{j}_n \times \vec{B}) \qquad (4)$$

$$\vec{j}_p = q(u_p p\vec{E} - D_p\nabla p) + u_p^*(\vec{j}_p \times \vec{B}) \qquad (5)$$

where u_n and u_p are electron and hole drift mobility respectively. D_n and D_p are the corresponding diffusion coefficient respectively. $u^* = 1.92u$ is the Hall mobility. Along the metal contact boundaries, the value of n, p and ψ are given by infinite recombination rate, charge neutrality and applied voltage V_A, that is

$$pn = n_i^2 \qquad (6)$$

$$p - n - N = 0 \qquad (7)$$

$$\psi = V_A + \frac{KT}{q}\sinh^{-1}\left(\frac{N}{2n_i}\right) \qquad (8)$$

Along the other insulated "floating" boundaries, the following condition is adopted.

$$\vec{j}_n \cdot \vec{r} = 0 \qquad (9)$$

$$\vec{j}_p \cdot \vec{r} = 0 \qquad (10)$$

$$\frac{\partial^2}{\partial r^2}(p - n + N) = 0 \qquad (11)$$

where $\vec{r}$ is the normal to the boundary. The partial differential equation (1)-(5) and the boundary condition (6)-(11) are discretized using the generalized Scharfetter-Gummel scheme discussed in [1]. Then, a decoupled iteration procedure is implemented, which treats the continuity equation and Poisson's equation separately.

2.2. Normal Region

In normal operation region, neglecting the hole current and electron diffusion current, the current continuity equation is converted into

$$\text{divgrad } U(x,y) = 0 \qquad (12)$$

where

$$U(x,y) = C_{ox}\Big(V_{FB} + \psi + \frac{C_{ox}}{2}\psi^2 + \frac{1}{3\varepsilon_{si}N_A q}(2\varepsilon_{si}N_A q\psi + 4\varepsilon_{si}N_A q\psi_B)^{3/2}\Big) \qquad (13)$$

here C_{ox} is the gate oxide capacitor

per unit area, ψ_B the Fermi potential, N_A the bulk doping density, V_{FB} the flatband voltage. At metal contact boundaries, U is known

$$U = U \mid \psi = V_A \qquad (14)$$

The current equation simply becomes

$$\vec{J} = \text{ugrad } U + u^* \vec{B} \times \vec{J} \qquad (15)$$

where J is the current density j_n integrated with respect to z over appropriate thickness.[2] At the insulated "floating" boundaries, the current normal to the surface must vanish, viz.

$$(\vec{r} + u^* \vec{B} \times \vec{r}) \text{ grad } U = 0 \qquad (16)$$

The partial differential equation and the corresponding boundary described above are discretized and numerically solved using Stone's iteration method.[3] Also, setting

$$V_G = V_D + V_{FB} + 2\psi_B + \frac{1}{C_{ox}} (2\varepsilon_{si} q N_A V_D + 4\varepsilon_{si} N_A q \psi_B)^{1/2} \qquad (17)$$

the numerical solution of MAGFET in linear region is extended to the saturation region.

3. RESULTS AN DISCUSSIONS

The parameter of fabricated MAGFET device is shown in the caption of Fig.3. The drain separation of d is 8 μm. The discretization is implemented on a non-uniform rectangular grid of 30 x30 nodes. As usual, the sensitivity S of a split-drain MAGFET device is defined by the relative current imbalance in the two drains $S = (I_1 - I_2)/(I_1 + I_2)$. The numerical results of S are plotted as a function of ratio (W/L) in Fig. 3. It indicates that the ratio (W/L) of 0.8 gives the highest sensitivity.

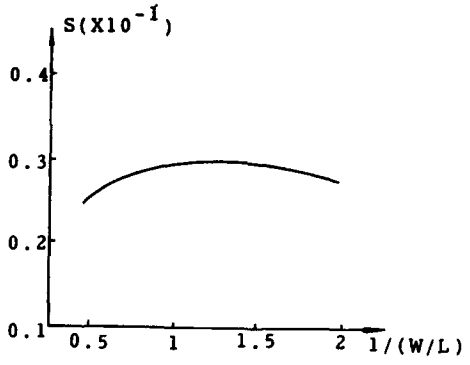


Fig.3 Sensitivity of MAGFET varies with (W/L)
W=28μm, t_{ox}=1200Å, B =1T, $N_A = 5X10^{15} cm^{-3}$, $V_{D1}=V_{D2}=1V$, $V_G=5V$

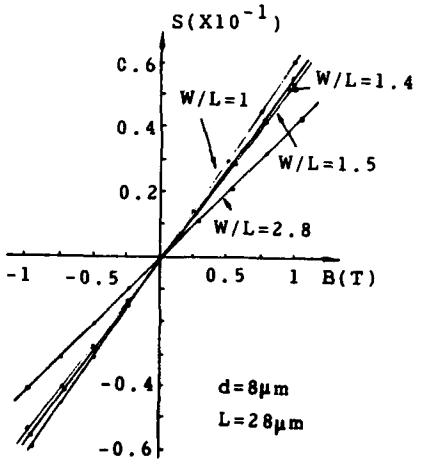


Fig.4 The device sensitivity varies with B for various aspect ratio W/L

As shown in Fig.4, the curves of measured device sensitivity versus magnetic field for various aspect ratio of W/L are plotted. Among the devices which aspect ratio is 1, 1.4, 1.5, 2.8, respectively, the aspect ratio near 0.8 (W/L=1) has the highest sensitivity. It supports the numerical results.

On the other hand, focusing attention

on the computation efficiency, the numerical modelling of MAGFET presented here is more attractive for a significant CPU time and computer memory saving capability. A set of typical curves of CPU time versus gate bias for various split-drain bias V_D is illustrated in Fig.5.

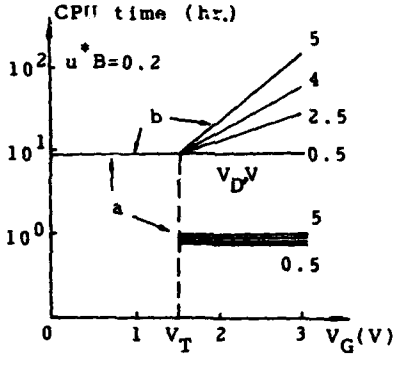


Fig.5 CPU time versus V_G for various V_D(on IBM PC/XT)
a. present method
b. decouped approach

4. CONCLUSIONS

In this paper, Magnetic-Field-Sensitive MOS device (split-drain MOSFET) is analyzed with computer. A two dimensional numerical simulation of the device is presented in subthreshold and normal region with different model equation. The computation cost is efficiently cut. The theoretical results show that the device of aspect ratio of 0.8 has the highest sensitivity. The experiment are in good support of that.

REFERENCES

1. Laszlo Andor et al, IEEE Trans. Electron Devices, vol.ED-32 (1985) 1224.
2. A. Nathan et al, IEEE Trans. Electron Devices, vol. ED-32 (1985) 1212.
3. H.L. Stone, SIAM J.Num. Anal., vol.5 (1968) 530

RESEARCH ON THE GROWTH CONDITION OF CdTe SINGLE CRYSTAL FOR γ-RAY DETECTOR

KATSUMI MOCHIZUKI and KATASHI MASUMOTO

Department of Materials Science, Faculty of Engineering,
Tohoku University, Aoba, Aramaki, Sendai, Japan

CdTe single crystals were grown by the Bridgman method from Te excess solution with a molar ratio of Cd/Te=3/7 and halogens were added to the solution for compensating acceptor center due to cadmium vacancies(V_{Cd}). Photoluminescence(PL) and electrical resistivity of the crystals grown under various conditions were measured for examining the incorporation of halogens and its compensation state. It was found that the halogen was incorporated in order from fluorine(F) to iodine(I). Halogen donors(D) incorporated under a growth rate of 6 mm/day and a temperature gradient($\Delta T/\Delta X$) of 20 K/cm form preferentially acceptor complex center(V_{Cd}·D) and this leads the crystal to p-type conductivity with unsatisfactory compensation. On the contrary, the donors incorporated under the growth rate of 72 mm/day (or steep temperature gradient) made the crystals into n-type high resistivity with enough compensation. Good γ-ray response was obtained only for the crystals with enough compensation. Therefore, relatively rapid growth or the growth under steep temperature gradient is required for realizing CdTe γ-ray detector.

1. INTRODUCTION

CdTe is known as one of the candidate materials for solid-state nuclear detectors operative at room-temperature[1,2]. So far, the main method to grow CdTe single crystal for γ-ray detector was from Te solvent which did extract most of impurities[1-4]. However, crystallization from an excess Te solution has also the disadvantage of introducing a native defect such as cadmium vacancies(V_{Cd}) in the grown crystals which reduce the electrical resistivity remarkably. As a very high resistance(10^7 ohm·cm) is usually required for CdTe single crystals used as a γ-ray detector [2,3], the above demerit has to be avoided by donor-doping for electrical compensation. Indium(In) and chlorine(Cl) have been used previously as compensation agents to obtain semi-insulating materials[3,4]. However, the possibility or usefulness of other halogens and the growth condition for obtaining optimum compensation state have not been discussed in detail. In this study, CdTe single crystals were grown by the Bridgman method and halogens such as F, Cl, I were added to the solution as a donor dopant for electrical compensation. The properties of doped crystals grown under various growth conditions were characterized by the measurements of photoluminescence(PL) and electrical resistivity and an indispensable growth condition for realizing CdTe γ-ray detector is discussed.

2. EXPERIMENTAL PROCEDURES

2.1. Growth of Halogen-doped CdTe Single Crystals

Cd with nominal purity of more than 6 nines and Te with that of 6 nines were used as the starting material for preparing CdTe single crystals. The elements with a molar ratio of Cd/Te=3/7 and their total moles of 0.17 were introduced into a carbon-coated quartz tube, together with one of the halogens of the amount 30 or 300 ppm per total weight and the tube was sealed at the vacuum of 1.3×10^{-4}Pa. The halogens were added in the form of cadmium halides such as CdF_2(5 nines), $CdCl_2$(4 nines) and CdI_2(4 nines). The sealed tube was put in a Bridgman furnace and the experiments for growing single crystals were performed under temperature gradient($\Delta T/\Delta X$) from 10 to 50 K/cm and the growth rate from 6 to 72

mm/day. Undoped CdTe single crystals were also grown from the melt with the composition of 5 % excess Te and from a Te excess solution with a molar ratio of Cd/Te=3/7.

2.2. Photoluminescence(PL) and Electrical Conductivity Measurements

Small samples about the size of 1x2x4 mm^3, which were cleaved from the grown crystals, were used for both experiments. PL spectra were measured at 4.2 K using He-Ne laser (632.8 nm) with 2 mW and 2 mφ beam as the exciting source. The detection system consists of a grating monochrometer(Nippon Bunko CT-50), a photomultiplier, a lock-in amplifier and a recorder. The spectral resolution of the luminescence is about 0.2 nm. The electrical conductivity was measured by a two-probe method. Au, decomposed from a chlorine solution, was deposited on the cleaved and unetched faces of the samples with a thickness of about 1 mm and was used as the electrodes.

3. RESULTS

Figure 1 shows PL spectra at 4.2 K of an undoped crystal grown from the melt with 5 % excess Te composition and of the crystals grown from a Te excess solution with a molar ratio of Cd/Te=3/7 and with halogens of 300 ppm per total weight. The growth experiments were performed under the growth rate of 6 to 8 mm/day and temperature gradient($\Delta T/\Delta X$)=20 K/cm.

It is known from figure 1(a) that the PL spectrum consists of band edge-emission such as A°X and D°X, weak B and C peaks near 1.55 eV and X peak series X-X'''' at about 1.45 eV. A°X (1.5896 eV) and D°X(1.5945 eV) are known as radiative emissions of excitons bound to a neutral acceptor and a donor respectively[2,5-8]. The peak similar to B(1.553 eV) and C(1.548 eV) have been reported[2,7] and these are probably due to radiative recombination of donor-acceptor(D-A) pairs. It is well-known especially for X peak(1.475 eV) that it is due to D-A pair between an isolated donor(D^+) and a complex acceptor($V_{Cd}^{--}\cdot D^+)^-$[9,10]. The notation with primes such as X'-X'''' shows phonon replicas of the X peak[6,11]. From figure 1, the following characteristics were observed; (1) In the case of F-doped, PL spectrum does not change much compared with that of an undoped crystal, though C and D peaks become obvious. (2) In the case of Cl and I-doped, excitonic emission A°X disappears and only the X peak is observed. The PL spectra shown in figure 1 are the cases of the crystals grown under a growth rate of 6 to 8 mm/day. On the other hand, a typical spectrum of the sample grown under relatively high growth rate such as 72 mm/day is shown in figure 2. This is the case of 300 ppm-doping of Cl. From the figure, excitonic emissions A°X and D°X were both observed with the same emission intensity and this is quite different from the spectrum for the same doping case shown in figure 1.

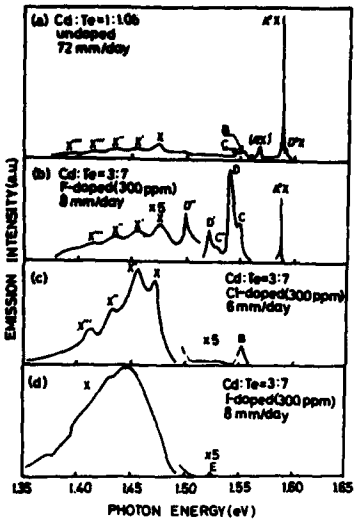


FIGURE 1

PL spectra at 4.2 K for undoped and halogen-doped CdTe

This result suggests that the growth rate strongly affects the crystal properties.

Figure 3 shows the electrical resistivity at room temperature of the samples doped with halogens. It is clear that the resistivity of the sample doped with F is almost the same as

that of an undoped one. The samples doped with Cl show the higher resistivity than others and the highest one is obtained for the sample grown under growth rate of 72 mm/day. The resistivity for the case of I-doped is lower than that of Cl-doped. The sample with the highest resistivity showed n-type, but the others were p-type.

Figure 4 shows the electrical conductivity of samples doped with Cl of 300 ppm as a

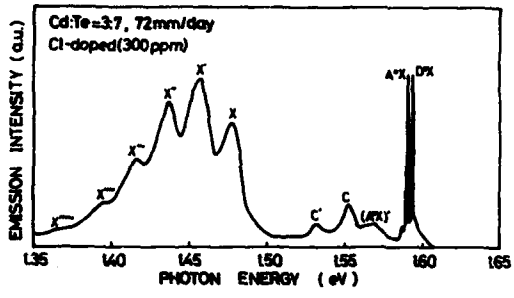


FIGURE 2

PL spectra at 4.2 K for Cl-doped CdTe crystal grown under growth rate of 72 mm/day

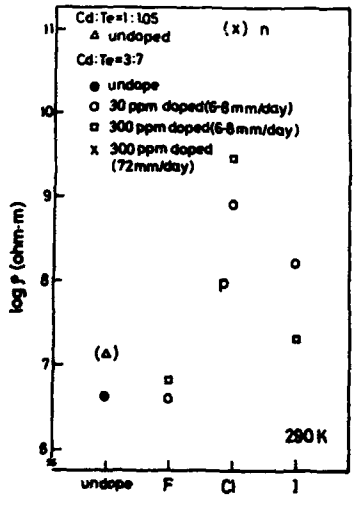


FIGURE 3

Electrical resistivities at room temperature of samples doped with halogens

function of temperatures. This suggests that deep levels which locate near the center of the band gap dominate the conductivity. In

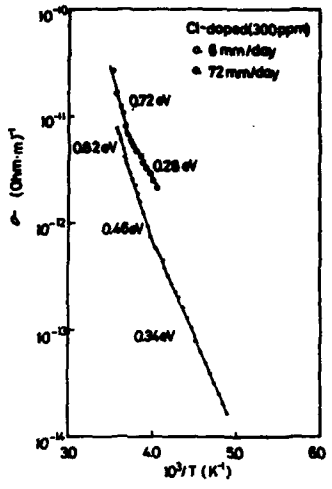


FIGURE 4

Electrical conductivities of samples grown under 300 ppm Cl-doped as a function of temperature

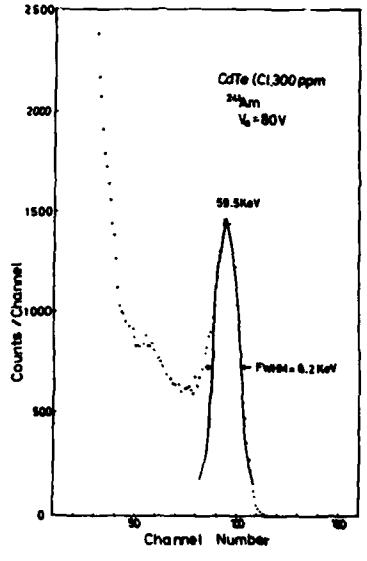


FIGURE 5

γ-ray spectrum of ^{241}Am source detected with CdTe(Cl,300 ppm doped) detector at room temperature

the case of 300 ppm F and I doped, shallow levels such as 0.18, 0.21, 0.28 and 0.37 eV were observed, suggesting unsatisfactory compensation. Even under a growth rate of 6 mm/day, extremely high resistivity crystals were also grown under high temperature gradient($\Delta T/\Delta X$).

Good γ-ray response was obtained only for the samples with extremely high resistivity due to enough compensation. Figure 5 shows a typical γ-ray spectrum of ^{241}Am obtained for such CdTe detector. A photopeak with FWHM of 8.2 keV was clearly observed. The solid line in the figure shows the peak shape calculated by assuming a Gaussian distribution. A tail is observed at the low energy side of the photopeak, suggesting hole trapping.

4. DISCUSSION

For the single crystals grown under growth rate of 6 to 8 mm/day and temperature gradient ($\Delta T/\Delta X$) of 10 to 20 K/cm, the incorporation and the behavior of halogens in the crystal are deduced as follows. In the case of F, PL spectrum and electrical resistivity are unchanged from those of an undoped crystal. This suggests that the incorporation of F is very little and should be the same order or less than the concentration of native defect such as V_{Cd}. On the contrary, in the cases of Cl and I, as seen in figure 1, PL spectra are quite different from those of an undoped one and it consists of only X peaks without excitonic emissions. This suggests that much quantities of Cl and I are incorporated in the grown crystal and they form a complex center $(V_{Cd}^{--}\cdot D^{+})^{-}$. The incorporated quantity can be estimated roughly from the similarity of PL spectra of In-doped CdTe with the well-defined amounts and it is 10^{16}, 10^{17} and $10^{18}/cm^3$ for 300 ppm-doping of F, Cl, I respectively[10]. The degree of the incorporation seems to depend on the atomic radius of halogens which are replaced on Te sites. The case of I has the closest radius to that of Te(Cd) and observed heavy incorporation of I supports this expectation. Much incorporation of the donor suggests the easy compensation of acceptor due to V_{Cd}. However, as seen in figure 2, the crystals show p-type conductivity with unsatisfactory compensation. This also supports that the halogens incorporated in the crystal preferentially form the acceptor complex center $(V_{Cd}^{--}\cdot D^{+})^{-}$. That is, the more halogens are incorporated, the more complex centers are formed. On the other hand, as seen in figure 2 and 3, the crystals grown under growth rate of 72 mm/day show n-type high resistivity and excitonic emissions recover. This suggests that an ideal compensation state is realized without forming preferentially a complex center. This should be the case in the situation that rapid growth prevents the incorporated halogens from diffusing to V_{Cd} site and increases the isolated donor concentration.

5. CONCLUSION

CdTe single crystals were grown by the Bridgman method from Te excess solution with halogens as a donor. The crystals with extremely high resistivity were grown under relatively high growth rate(or high temperature gradient) and it was concluded that the growth condition is indispensable for realizing γ-ray detector.

REFERENCES

[1] Bell, R.O.,Hemat,N. and Wald,F., Phys. Stat. Solidi 1(a)(1970)375.
[2] Triboulet, R.,Marfaing,Y.,Cornet,A. and Siffert,P., J.Appl.Phys.45(1974)2759.
[3] Hage-Ali,M., Scharager,C.,Koebel,J.M. and Siffert,P., Nucl. Instr. Methods 176(1980) 499.
[4] Höschl,P.,Polívka,P.,Prosser,V.,Vaněček,M. and Skřivánková,M., Rev.Phys.Appl.12 (1977)229.
[5] Taguchi,T.,Shirafuji,J. and Inuishi,Y., Phys.Stat.Solidi(b) 68(1975)727.
[6] Molva,E.,Chamonal,J.P. and Pautrat, J.L., Phys.Stat.Solidi (b) 109(1982)635.
[7] Francou,J.M.,Saminadayar,K.,Pautrat,J.L., Gaillard,J.P.,Million,A. and Fontaine,C., J.Cryst.Growth 72(1985)220.
[8] Triboulet,R.,Legros,R.,Heurtel,A.,Sieber, B.,Diedier,G. and Imhoff,D., J.Cryst. Growth 72(1985)90.
[9] Agrinskaya,N.V., Arkadeva,E.N., and Matveev., Sov.Phys.Semicond.5(1971)767.
[10] Barnes,C.E. and Zanio,K., J.Appl.Phys.46 (1975)3959.
[11] Halsted,R.E., Lorenz,M.R. and Segall,B., J.Phys.Chem.Solids 22(1961)109.

TRENDS IN THREE-DIMENSIONAL INTEGRATION

Yoichi Akasaka

Mitsubishi Electric Corp., LSI R&D Laboratory,
4-1 Mizuhara, Itami,
664, Japan

VLSI will be reaching to the limit of minimization in the 1990s, and after that, further increase of packing density or functions might depend on the vertical integration technology and wafer scale integration. The 3-D ICs consisting of completely stacked active layers offers the flexibility of circuit design and composition of various devices. Three-dimensional (3-D) integration is expected to provide several advantages, such as 1) parallel processing, 2) high-speed operation, 3) high packing density, and 4) multifunctional operation.
This will lead upto new system design and the novel functional device. It will become a big trend for VLSI in the next generation.

INTRODUCTION

The ultimate IC structure of the future is thought to consist of stacked active IC layers sandwiched by insulating materials or wafer sale integration.

Various devices or circuit functions, such as photosensors, logic circuits, memories, and CPUs, will be arranged in each active layer and, as a result, a remarkable improvement in packing density and functional performance will be realized.

In 1979, it was reported that polysilicon deposited on insulator can be melted and recrystalized by laser irradiation [1] and that the crystal perfection of the layer can be adequate to allow the fabrication of devices.

The quality of the recrystallized layer can be characterized by carrier mobility. The electron mobility reported so far has increased year by year and has attained a value comparable to bulk crystals. This improvement was a trigger for starting research and development of 3-D ICs.

A partial 3-D structure has already been tried for a dynamic memory (DRAM) cell [2 , 3], For high-density RAMs, such as the 4-M bit DRAM, the area of the memory cell capacitor is limited, so in order to increase the cell capacitance, a 3-D structure, i.e., a trench cell or a stacked capacitor cell, has been tried. This partial 3-D structure will be used in 2-D VLSIs within a few years.

To achieve a breakthrough in the packing density of advanced VLSIs, it is reasonable to consider a 3-D structure, containing either partially or completely stacked active layers.

Basic technologies of 3-D IC are to fabricate SOI layers and to stack them monolithically. Crystalinity of the recrystallized layer in SOI has increasingly become better, and very recently crystal axis controlled, defect-free single-crystal area has been obtained in chip size level by laser recrystallization technology.

Some basic functional medels showing the concept or image of a future 3-D IC were fabricated in two or three stacked active layers.

Some other proposals of subsystems in the application of 3-D structure, and the technical issues for realizing practical 3-D IC, i.e., the technology for fabricating high-quality SOI crystal on complicated surface topology, crosstalk of the signals between the stacked layers, total power consumption and cooling of the chip, will also be discussed in this paper.

3-D IC STRUCTURE

Fig. 1 shows a typical basic structure of

3-D devices proposed by several researchers. Fig.1(a) is a flip-chip, which is an attempt to realize the stacked 3-D form by chip assembly technology. This technology has already been used in computers as a connection between a group of single chips and a printed circuit board. In this technology, the number of connections are restricted by reliability and bump size constraints. Fig.1(b) shows a new approach based on wafer process technology. The chips are attached face-to-face by pressure [4]. The minimum connection area is 10μm^2, which is large compared with the device feature size (1-2 μm), but the number of connections will be greatly increased by this technology. Fig.1(c) and (d) shows a monolithic 3-D structure. At first, the passive elements, such as the DRAM-cell capacitor or the load resistor of a SRAM, will be stacked three-dimensionally similarly to the multilevel interconnections of today's LSIs. Fig.1(c) shows 3-D integration performed at the transistor level [5]. Load transistors of a CMOS inverter of a CMOS static RAM cell are fabricated in the upper layer to form a partial 3-D structure. The first 3-D IC which has active elements or uses single-crystal material in the stacked layer may have the configuration shown in Fig. 1(c). Fig.1(d) shows the stacked form of LSI layers in a complete 3-D structure, which is also a final goal of the 3-D Project in Japan. In this structure, the degree of freedom in circuit or system design circuit layout, and the reliability of the interconnections are expected to be very high. This paper will discuss mainly the trends in the 3-D structure shown in Fig.1(d).

FEATURE OF 3-D IC

The advantageous features expected for 3-D IC as shown in Fig.1(d) are follows [6];

(1) High packing density, or super large integration becomes possible without having a serious problem about power dissipation. Because the multi-layer devices in 3-D IC can have one set of I/O circuits in common.

(2) High speed performance of 3-D ICs is associated with shorter interconnection delay time and decrease of parasitic capacitance due to the use of SOI structure.

(3) Parallel processing is one of the special features of the 3-D IC. Widely different large number of information signals could be transferred from upper to bottom layers, or vice-versa through via-holes with 1-2 μm diameter.

(4) Multi-function is also one of the special features of the 3-D IC. Each layer, or a set of several layers can have their own functional performance. Also at the device level, different performances, MOS and bipolar, and fifferent characteristics caused by different process technologies can be assigned to each desired active layer. It will be possible to make use of these advantages for system design.

BASIC TECHNOLOGIES

(1) SOI Technology

The basic technology for 3-D ICs is to form a silicon-on-insulating (SOI) layer, and to stack it monolithicaly. Recrystallization of a polysilicon film by laser- or electron-beam is thought to be the most suitable method because the wafer temperature during crystallization is kept low, and this leads to the good stability of the underlying device performance.

The basic concept in the laser-recrystallization [7] of the partially restricted active area is to control the temperature profile in the polysilicon layer like twin peaks or periodic profile [8, 9] as shown in Fig. 2. The crystallization proceeds from one nuclei of polysilicon along the desired direction.

Fig.3 summarizes specific methods of realizing the preferable thermal profile in the molten zone of the polysilicon:

a) shaping the intensity profile in the laser or electron beam itself;
b) obtaining a thermal profile by changing the energy absorption characteristics by pat-

terning the antireflecting thin film or absorption layer on the top of the polysilicon;

c) changing the heat transfer selectively by the material design of the sample structure.

Very recently, the fundamental relation between the crystal quality of laser-recrystallized SOI and the crystallographic arrangement of growth front was studied [10], and thereby the new technique for obtaining large single crystal SOI. Figure 4 shows the length of the single crystal regions from the seed edge versus laser scan direction. The schematic illustration of the sample structure and its crystallographic arrangement based on a (001) Si seed is also shown in Fig.4, which is the combination of the lateral seeding technique [11] and use of the patterned antireflecting film [8] to control the shapes of the liquid-solid interfaces. In a <100> direction, single crystal growth has reached to the chip end without any defects except the controlled subgrainboundaries beneath the antireflecting stipes.

Electrical characteristics of MOS devices on this SOI film and their uniformity through the wafer were extensively improved. The divergence of the threshold voltage of MOSFET/SOI was reduced as below 8 %, and was successfully applied for fabricating a 4 bit CMOS A/D converter which was the first analogue devices formed on SOI [12, 13].

(2) Planarization

Insulator surface of each active layer in the 3-D structure should be planarized less than +0.1 um accuracy for obtaining the high quality SOI film on it. Reflow of silicate glass (PSG or BPSG), sputter etching, [14] spin-on-glass coating and etching back technique with organic resist material [15] have been investigated.

(3) Interconnection Material

The interconnection material should be resistive to the heat cycle subjected in the fabrication process steps for upper active layers. Refractive metal or its silicide [13] and phosphorus doped polysilicon [9] were examined.

(4) Via-Holes

Via-holes with very high aspect ratio (i.e. 1 um diameter hole with the height of 2-3 um) will be required. In addition to fine patterning and etching, selective growth or deposition of conductive materials in a via-hole is required.

(5) Low Temperature Process

Process temperature should be lowered to avoid the redistribution of doping profiles in already finished devices in the bottom layer.

MULTI-LAYER STRUCTURE

Up to the present, a few kinds of 3 levels stacked structure have been reported. Figure 5 shows the SEM cross sectional view and a schematic drawing of 3-level 3-D IC (2 levels of SOI and one bulk-Si) [16]. Each layer has its own interconnection and can be operated independently, and also electrically connected vertically with another layer through via-holes. Average surface electron mobilities obtained from 2nd and 3rd layers were close to that of bulk-Si devices.

3-D DEVICE

Several small-scale test devices have been fabricated to make clear the concept and the future image of the 3-D IC. The image sensor with a proper level of intelligence is one of the important target of 3-D ICs, because it can make use of the feature of 3-D integration and it will have extremely wide application area.

Based on state-of-art 3-D device technologies and on the concept of 3-D IC, the new idea of the real time image processor was pro-

posed as shown in Fig 6 [17]. The 6 layered 3-D LSI consists of 512-by-512 optical sensor, A/D converter, memory, switch matrix, accumulator, and memory. The main function is the feature extraction, in which a spatial convolution of a 3-by-3 pixel kernel are performed for 512-by-512 pixels. The system speed is estimated to become 10000 times over a conventional serial processing by typical super-minicomputer.

Figure 7 shows one of the small-scall test devices for image signal processing fabricated in 3-stacked ICs. The first top layer contains 5x5 image sensor pixcells. The 25 2-bit A/D converters which are connected to corresponding 25 pixcells are fabricated in the 2nd layer. This enables the parallel processing of the image signals. The 3rd layer consists of 32 kinds of ALU, each of which acts as an image signal processing, as feature extraction, noise reduction or image enhancement etc. The performance of this system through 3 levels of IC is shown in Fig. 8.

In addition to those synthetic function, there are some other proposals of application with new concepts [18]. Those are a PLA by a simple combination of AND- and OR-block prepared in top and bottom layers, respectively, a Logic-in-Memory which assign the two-dimensionally accessed memory for video processing on the top layer and the corresponding microprocessor on the bottom layer, and a CAM which has the same configuration with logic-in-memory, etc.

PROBLEMS

The problems which might be serious in future 3-D devices were studied [6]. Limiting factor of total power dissipation and heating of a chip is mainly the thermal resistivity from inside the package to the outside in 3-D structure as well as in 2-D VLSI. Therefore the problems of heating in 3-D IC are essentially the same as with the case of 2-D VLSI, as far as the integration level of devices is the same. It will become a problem in ultra large scale integration. For example, the device containing 10 million elements, if each element dissipates 5 uW, dissipates 50 W. Therefore, an innovation of the packing technology will be required in future.

Signal cross talk will be more serious in 3-D IC than 2-D VLSI, which includes the cases of DC back gate bias to back channel of SOI transistors and high frequency operation of the circuits [19]. For high frequency operation, it has been investigated that signals of two circuits arranged in a stacked from with upper and lower layers affected each other, in which one signal was superposed to another signals as a noise when two frequencies were far apart, and two signals were locked in the same frequency when two frequencies became close. It has been also reported that the cross talk depends on the structural parameters as well as the layout and the performance of the devices. A shielding plate between upper and lower devices may be necessary for preventing the cross talk in 3-D structure.

SUMMARY

The 3-D IC is now in the early stage of fundamental research. Extensive work is still required in both process technology and architecture design of chip and system for realizing the sophisticated 3-D IC. However, memory or logic chip which has even 2 stacked layers is worthy for commercial use if it is produced at reasonable cost under the stable and reliable fabrication technology. This may enhance the 3-D IC development and push forward the time table for wide spread availability of 3-D ICs in more complicated multi-layer structure.

ACKNOWLEDGEMENT

Most data shown in this paper were obtained in the work of 3-D project under the management of the R&D Association for Future Electron Devices as a part of the R&D Project of

Technology for Future Industries sponsored by agency of Industrial Science and Technology, MITI.

The authors would like to express their gratitude to Drs. K. Shibayama and H. Nakata for encouragement.

REFERENCES

(1) A. Gat, L. Gerzberg, J.F. Gibbons, T.J. Magee, J. Peng, and FLD. Hong, Appl. Phys. Lett., vol. 41, p. 775, 1978.

(2) K. Itoh, R. Hori, J. Etoh, S. Asai, N. Hashimoto, K. Yagi, and H. Sunami, in 1984 ISSCC Tech. Dig., paper FAM 18.6, p. 282, Feb. 1984.

(3) M. Taguchi, S. Ando, S. Hijiya, T. Nakamura, S. Enomoto, and T. Yabu, in 1984 ISSCC Tech. Dig., paper WPM8A-3, p. 100, Feb. 1984.

(4) M. Yasumoto, H. Hayama, and T. Enomoto, in 1984 IEDM Tech. Dig., paper 34.7, p. 816, Dec. 1984.

(5) S.D.S. Malhi et al., IEEE Trans. Electron Devices, vol. ED-32, p. 258, J.P. Colinge, E. Demoulin, and M. Lobet, IEEE Trans. Electron Devices, vol. ED-29, p. 585, 1982.

(6) Y. Akasaka; Proceedings of The IEEE 74(1986)1703

(7) D.K. Biegelsen, etal; Appl. Phys. Lett., 38(1981)150

(8) J.P. Colinge et al; Appl. Phys. Lett., 41(1982)346

(9) T. Nishimura et al; Extended Abstract, 16th(1984 International) Conf. Solid State Devices and Materials p. 527

(10) K. Sugahara et al; Extended Abstract, 18tg(1986 International) Conf. Solid State Devices and Materials p. 565

(11) H.W. Lam et al; J. Electrochem. Soc., 128 (1983)1981

(12) S. Kusunoki et al; 1987 Symp. VLSI Tech. X-8, p107 (Karuizawa,1987)

(13) T. Kumamoto et al; 1987 Symp. VLSI Circuit, p111 (Karuizawa, 1987)

(14) M. Morimoto et al; 1983 Symp. VLSI Tech. 7-8 p.100

(15) . Mitsuhashi; 4th FED Symp. p.251 (1985, July)

(16) S. Sugahara et al; IEEE Elec. Dev. Let., EDL-7(1986)193

(17) K. Taniguchi; IEEE 7th Symp. on Computer Arithmetic, June 1985

(18) Reviewed by T. Ohtsuki, 4th FED Symp. V-1 (1985, July)

(19) Y. Akasaka et al; 43rd Dev. Res. Conf. VB-7 (1985, Corolado)

3-D STRUCTURE

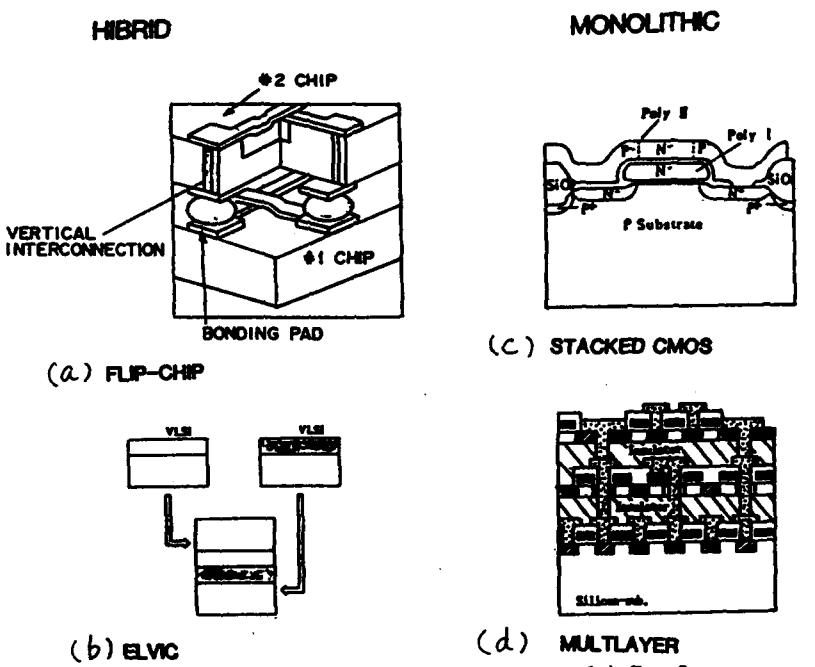
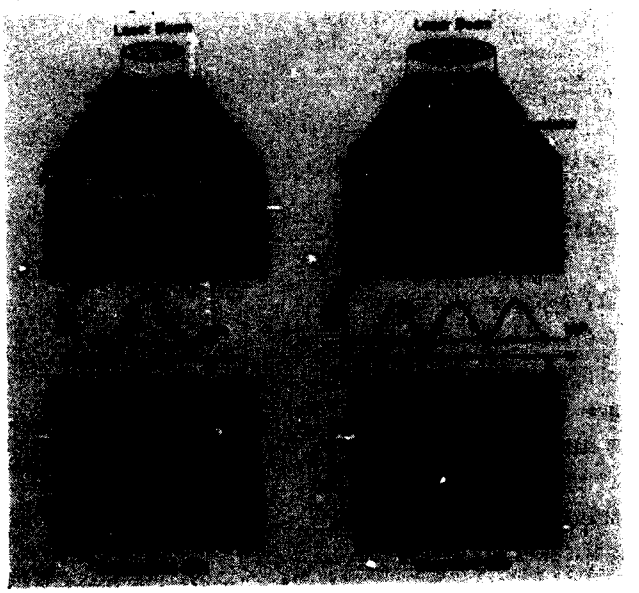


Fig. 1 Basic Concepts of 3-D device structure (a) Two 2-D LSI connected by flip-chip bonding. (b) Chip atta chiment by press. (c) Partially stacked structure in transistor level in monolithic 3-D IC. (d) completely stacked LSI in 3-D monolithic structure.



Selective Laser Recrystallization Technique

Fig.3 Periodic thermal profile caused by antireflecting stripes arranged onto the polysilicon layer.

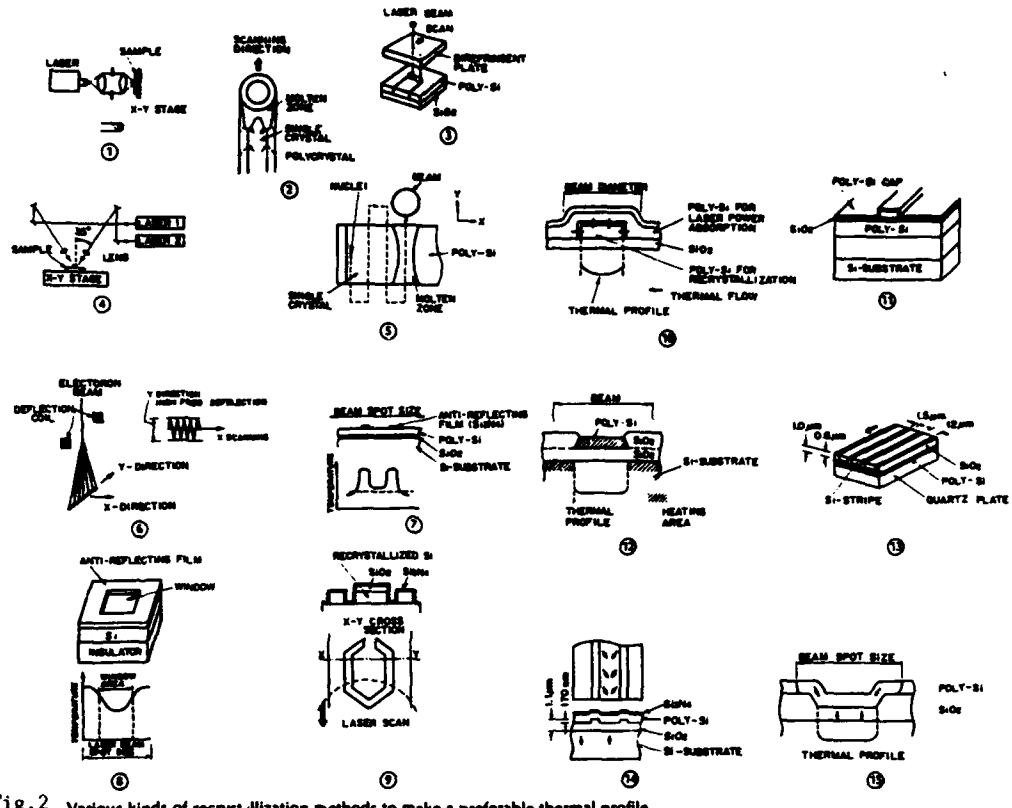
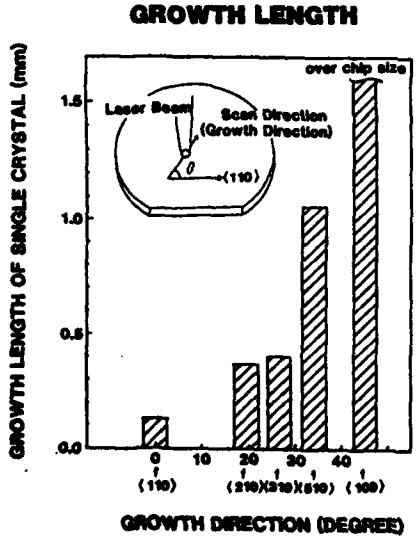


Fig. 2 Various kinds of recrystallization methods to make a preferable thermal profile in the molten zone of the polysilicon. a) By the intensity profile of laser or electron-beam spot ① Beam profile change by using mask (Stanford University, single crystal of 45 μm × 50 μm). ② Donut-type beam by oscillation mode modification (Fujitsu, crystal of 600-nm length). ③ Beam-splitting type (NEC, crystallized area of 20 μm × 1 mm). ④ Double laser beams (Fujitsu, 20 μm; Matsushita, 1.8 mm). ⑤ Electron-beam oscillation—osciliatory growth method (AT&T Bell Labs., 50 μm × 50 μm). ⑥ Quasi-linear electron beam (Toshiba, Tokyo Institute of Technology, 300 μm). b) By patterning the antireflection thin film or absorption layer on the top of the polysilicon (selective recrystallization). ⑦ Stripe-patterned antireflecting thin film (CNET, Mitsubishi, 20 μm × 400 μm). ⑧ Patterned antireflecting thin film (Fujitsu). ⑨ Changing the reflectivity (moated island) (General Electric, 18 μm × 50 μm). ⑩ Indirect heating (Fujitsu, 20 μm × 60 μm). ⑪ Double poly-Si layer recrystallization (Sharp). c) By modifying the heat transfer from molten zone of poly-Si. ⑫ Locos island (edge heating (TI, HP, Mitsubishi, Matsushita, Sharp, 10 μm × 50 μm). ⑬ Buried-stripe structure (NEC, 12 μm × 500 μm). ⑭ Relief structure of SiO_2 (control of thermal flow to a substrate) (Mitsubishi, 8 μm × 200 μm). ⑮ Heat sink structure (Fujitsu).



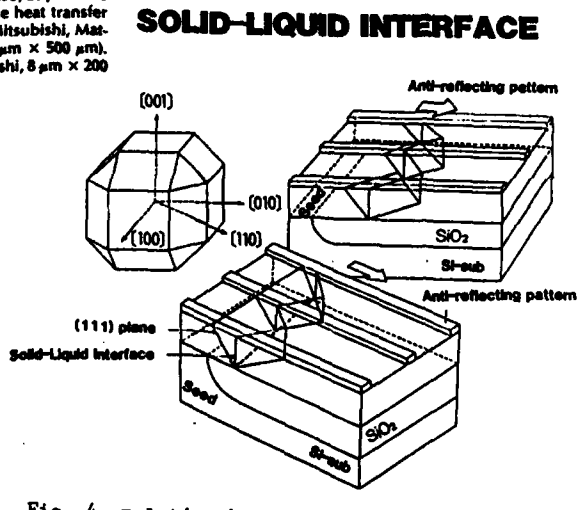


Fig. 4 Relation between the growth length of the single crystal from seed edge and crystallographic arrangement (growth direction)



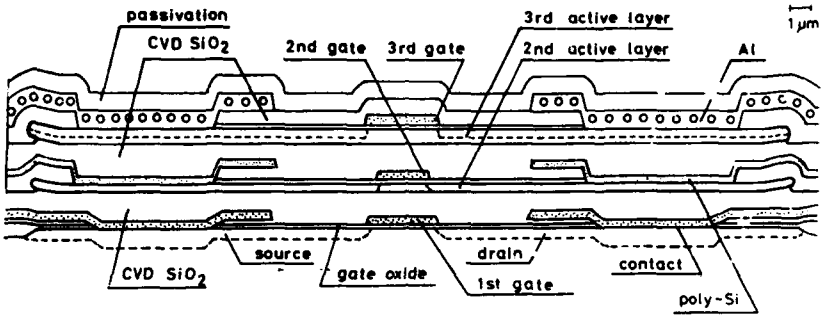


Fig. 5 SEM cross sectional photograph and schematic drawing of planarized 3 level 3-D IC structure.

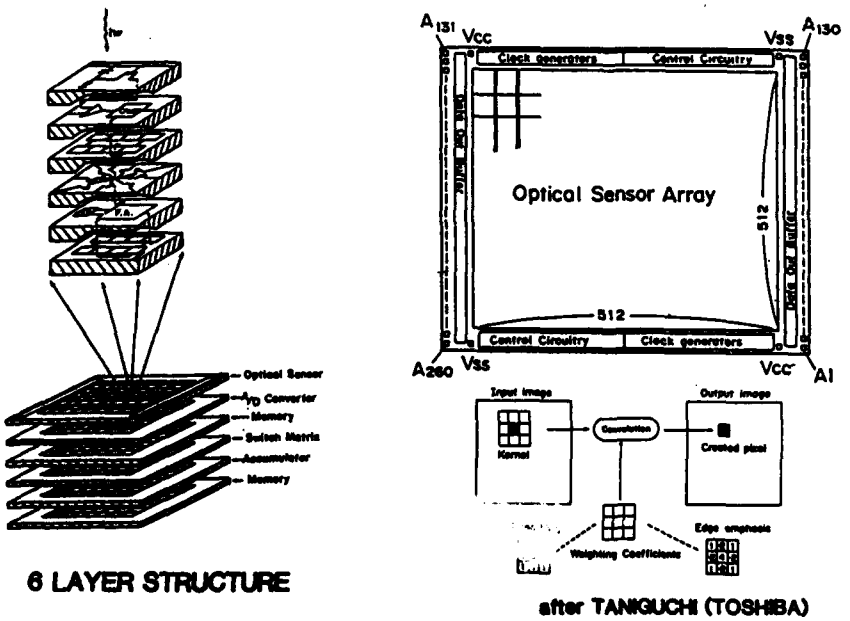


Fig. 6 Proposal of the real time image processor by 6 layered 3-D LSI (after Taniguchi)

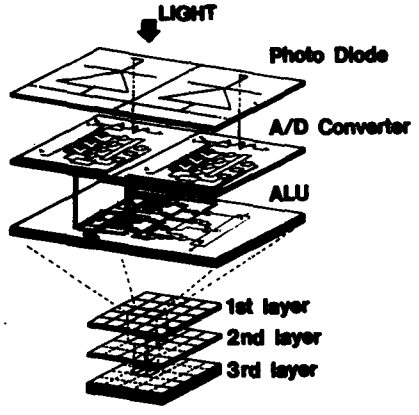


Fig. 7 Test device of one chip image processor fabricated in 3 layered 3-D IC.

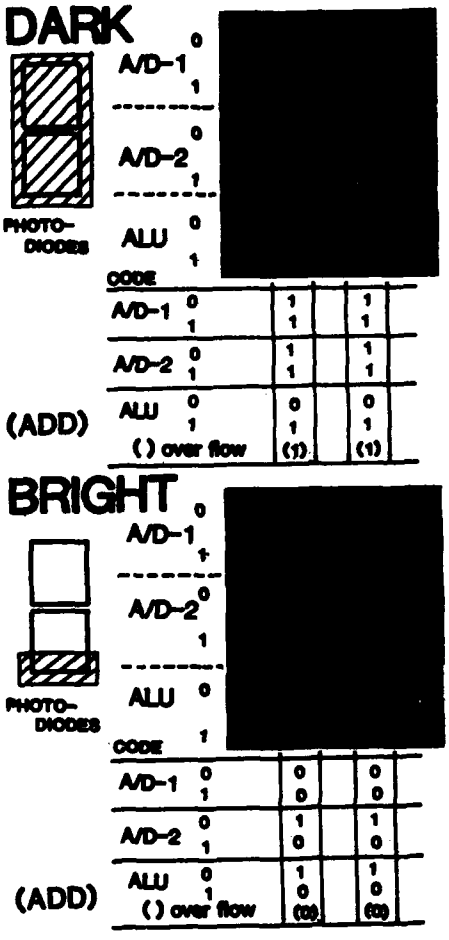


Fig. 8 Synthetic operation of the 3-D image processor measured from a pair of adjacent pixels. One pixel is shielded 1/3 of its surface by Al layer to control the amount of incident light.

Session A2.1

Ultrafast Bipolar I

Chairman: L. Treitinger

Tuesday, September 15, 1987

HISTORY, PRESENT TRENDS, AND SCALING OF SILICON BIPOLAR TECHNOLOGY

Tak H. Ning

IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598
USA

Recent advances in bipolar technology are reviewed. The key features of the advanced bipolar devices are identified and the trends for future development discussed. The scaling of high-speed circuits and the properties of the scaled devices are also reviewed.

1. INTRODUCTION

There is no question that advances in CMOS technology have generated most of the excitements in silicon VLSI technology in recent years. Perhaps more surprizing to most people is the fact that there have been very exciting development in the silicon bipolar front as well. Not only has bipolar technology remained the technology of choice for applications where circuit speed is of primary importance, it has also demonstrated its VLSI capability as well. The last few years saw particularly exciting developments in silicon bipolar technology, including sub-50ps NTL and ECL [1,2], high-speed as well as high-density RAM chips [3-5], 8b A/D converter in the 350-400MHz range [6,7], and frequency divider approaching 10GHz [8].

In this paper, the breakthrough responsible for this sudden acceleration in bipolar technology development is discussed. By examining the key elements of the advanced bipolar technology, the directions and trends for future development are established. The properties of the scaled bipolar devies are also discussed.

2. ADVANCED BIPOLAR TECHNOLOGY

There have been many reports on novel bipolar device structures and/or processes, such as PSA and APSA [9], SST [10], SICOS [11] or symmetrical transistor [12], BEST [13], SCOT [14], OXIS-III [15], and self-aligned bipolar [16]. Figure 1 shows the schematics of two of these devices together with deep-trench isolation [17,18]. It clearly illustrates the three key features of the advanced bipolar technology, namely (i) self-aligned base contact, in this case using polysilicon, (ii) deep-trench isolation, and (iii) polysilicon emitter contact. Both the self-aligned structure and the trench isolation greatly reduce the device area and the associated parasitic capacitance, and hence significantly reduce the power-delay product and increase the density of bipolar circuits. The various advanced devices and processes cited above represent efforts by different companies to develop their own version of self-aligned device structure. Most self-aligned device structures have comparable but low parasitic capacitance, at least as far as high-speed circuit applications are concern. The focus of future development effort will be on reducing process complexity, compatability with narrow-base process, and extendability to submicron dimensions.

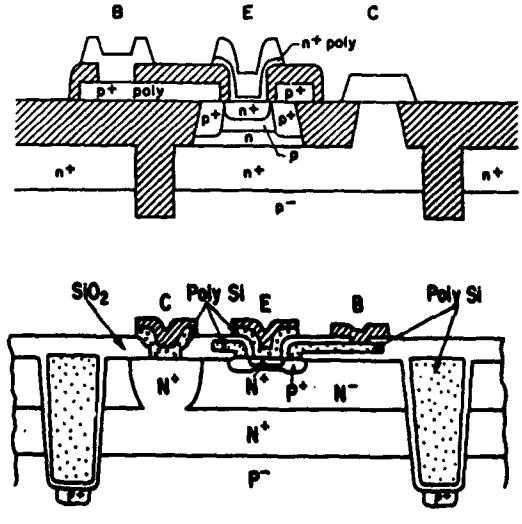


Fig. 1. Schematics of two advanced bipolar transistors, from Ref. [17] (top) and Ref. [18] (bottom), showing their three key features, namely i) self-aligned base contact, ii) deep-trench isolation, and iii) polysilicon emitter contact.

Deep-trench isolation has the biggest leverage in memory applications. That is why the early applications of deep-trench isolation were mostly to memory designs [19,20]. Of course, deep-trench isolation does improve logic circuit speeds as well [17,20]. The trenches have been filled with oxide [17], or with oxide and/or nitride and then filled with polysilicon [20,21]. Conceptually, the polysilicon-filled trenches cannot be made narrower than the oxide-filled trenches due to the thickness of the liners. However, as long as the trenches are not too wide, the leverage of an extra-narrow trench is not clear. The trend is therefore to reduce process complexity and not necessarily to reduce trench width.

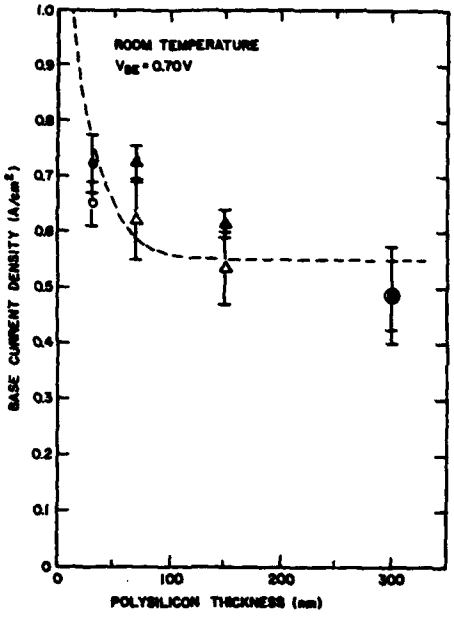


Fig. 2. Base current density as a function of the thickness of the emitter polysilicon contact layer [22]. Data represented by the same symbol were from the same wafer. It suggests that the polysilicon should be thicker than about 50 nm for maximum current gain.

The polysilicon emitter contact by itself has negligible effect on circuit speed [16]. Its primary importance is to improve the current gain so that the bipolar device can be scaled down vertically without having emitter-collector punchthrough problem. The data [22] shown in Fig. 2 suggest that the emitter polysilicon layer thickness should be greater than about 50 nm. Once current gain is adequate, the most important factor in determining the polysilicon emitter process is the emitter series resistance, usually attributable to the polysilicon-silicon interface. Significant circuit speed degradation can result if emitter resistance is not controllably low [23].

3. BIPOLAR SCALING

The exponential I-V relationship gives bipolar devices high transconductance. It also implies that the diode turn-on voltage remains relatively constant in scaling, increasing at a slow rate of about 60mV/decade increase in current density. As a result, the power-supply voltage remains approximately constant in bipolar scaling.

The design and scaling procedure for high-speed ECL circuits have been discussed in detailed in [24,25]. The central idea in ECL scaling, which is applicable to bipolar circuits in general, is to reduce the horizontal and the vertical dimensions in a coordinated manner so that all the key delay components are reduced approximately proportionately in scaling. The scaling relues for ECL circuits are shown in Table 1. It suggests that circuit delays in proportion to the emitter-stripe width can be expected. Figure 3 is a plot of the ECL and NTL ring-oscillator circuit speeds as a function of the emitter-stripe width, compiled from several recent publications. Although there is quite a bit of scattering in the data, it does indicate that the advanced self-aligned bipolar devices can be scaled down to sub-micron dimensions and achieve performance improvements. Of course, consistent with the concept of scaling, the other horizontal and vertical dimensions should be reduced with the emitter-stripe width in order to achieve the optimal design.

Table 1. ECL Scaling Rules [25] (a: lithography dimension and emitter width)

Parameter	Rule
Base width, W_b	$a^{0.8}$
Base doping level, N_b	W_b^{-2}
Collector current density, J_c	a^{-2}
Collector doping level, N_c	J_c
Circuit delay	a
Supply voltage	constant

The scaling rules in Table 1 also suggests that in order to achieve scaled circuit delay, the current, and hence the power dissipation, of the circuit remains essentially constant in scal-

ing. Thus, with power dissipation per circuit remaining in the milliwatt range for high-speed designs, the integration level of scaled ECL logic chips will be severely limited by the total chip power dissipation. Circuit and design innovations are needed to overcome this limitation and extend the application of bipolar technology to VLSI integration level [26].

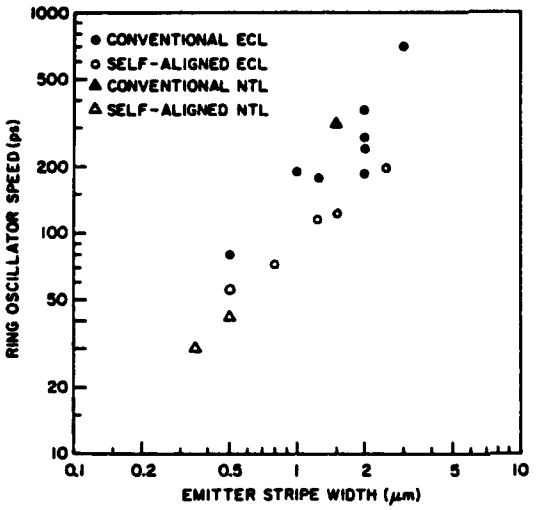


Fig. 3. Bipolar ECL and NTL ring-oscillator speed as a function of emitter-stripe width, compiled from recent publications. It show that the self-aligned devices are much faster than the conventional non-self-aligned ones, and can achieve speed improvement in scaling down to sub-0.5μm dimensions.

4. PROPERTIES OF SCALED BIPOLAR DEVICES

The physics of the advanced bipolar devices, particularly in the submicron dimensions, remains to be studied extensiveiy. An in-depth understanding of the physics of the submicron devices is needed in order to realize the performance potential suggested by the scaling theory. However, many of the properties of the scaled bipolar devices can be anticipated from examining the scaling rules in Table 1.

The combination of constant supply voltage and increased base doping level suggests that excessive leakage current due to tunneling [27,28] and hot-carrier degradation [29] in the emitter-base junction could become problems in scaled devices. Device and circuit design constraints due to these effects remain to be established.

While heavy-doping or band-gap narrowing effect in the emitter region degrades the current gain by enhancing the base current, heavy-doping effect in the base region increases the current gain by enhancing the collector current. Thus, with the base doping level increasing in scaling, current gain is not expected to be a problem in scaled devices [30]. Furthermore, as the base band-gap is reduced due to heavy-doping effect, the collector current becomes less sensitive to temperature variation. It should therefore be possible to achieve adequate current gain at temperatures as low as 77K for properly designed scaled devices [30,31].

The rapid increase in current density implies that current-stress-induced device instability could also occur in scaled devices [32].

There is always a perimeter transistor, associated with the perimeter emitter-base junction, in parallel with the intrinsic transistor, associated with the center or areal emitter-base junction. As the eimtter-stripe width is scaled down to submicron dimensions, the perimeter could become more and more dominant unless special effort is made to minimize its effect. Thus, electrical characteristics of the submicron device could become a function of the emitter-stripe width, giving rise to the so-called narrow-emitter effect. This effect is a strong function of the details of the device structure as well as the details of the emitter and base formation processes. Just as short-channel effect must be considered in submicron MOS device design, narrow-emitter effect must be considered in submicron bipolar device design [33-35]. Significant performance degradation could result if design optimization does not include narrow-emitter effects.

5. SUMMARY

Today's advanced bipolar transistors are generally characterized by one or combinations of the following three salient features, namely (i) self-aligned structure, (ii) deep-trench isolation, and (iii) polysilicon emitter contact. High-speed ECL or ECL-like circuits can have delays reduced approximately linearly with dimension in scaling, but only at approximately constant power dissipation. Many new device physics can be anticipated from the bipolar scaling theory. They must be understood in order to realize the full potential of the scaled advanced bipolar technology.

REFERENCES

[1] S. Konaka et al., "A 30ps Si bipolar IC using super-self-aligned process technology," in Extended Abstracts 16th Int. Conf. Solid-State Devices and Materials, pp. 209-212, 1984.

K. Washio et al., "A 48ps ECL in a self-aligned bipolar technology," in ISSCC Dig. Tech. Papers, pp. 58-59, 1987.

[3] H. Miyanaga et al., "A 0.85ns 1Kb bipolar ECL RAM," in Extended Abstracts 16th Int. Conf. Solid-State Devices and Materials, pp. 225-228, 1984.

[4] C.T. Chuang et al., "A 1.0ns 5Kb ECL RAM," in Extended Abstracts 18th Int. Conf. Solid-State Devices and Materials, pp. 267-270, 1986.

[5] T. Awaya et al., "A 5ns access time 64Kb ECL RAM," in ISSCC Dig. Tech. Papers, pp. 130-131, 1987.

[6] Y. Yoshii et al., "An 8b 350MHz flash ADC," in ISSCC Dig. Tech. Papers, pp. 96-97, 1987.

[7] Y. Akazawa, "A 400MHz 8b flash A/D Conversion LSI," in ISSCC Dig. Tech. Papers, pp. 98-99, 1987.

[8] M. Suzuki et al., "A 9-GHz frequency divider using Si bipolar super-self-aligned process technology," IEEE Electron Device Lett., vol. EDL-6, pp. 181-183, 1985.

[9] H. Nakashiba et al., "An advanced PSA technology for high-speed bipolar LSI," IEEE J. Solid-State Circuits, vol. SC-15, pp. 455-459, 1980.

[10] T. Sakai et al., "Gigabit logic bipolar technology," Electron. Lett., vol. 19, pp. 283-284, 1983.

[11] T. Nakamura et al., "Self-aligned transistor with sidewall base electrode," in ISSCC Dig. Tech. Papers, pp. 214-215, 1981.

[12] D.D. Tang et al., "A symmetrical bipolar transistor," in IEDM Tech. Dig., pp. 58-61, 1980.

[13] M. Shimuzu and H. Kitabayashi, "BEST (base-emitter self-aligned technology), a new fabrication method for bipolar LSI," in IEDM Tech. Dig., pp. 332-335, 1979.

[14] T. Hirao et al., "A 2.1-GHz 56-mW two-modulus prescalar IC using salicide base contact technology," in Extended Abstract 17th Conf. Solid-State Devices and Materials, pp. 381-384, 1985.

[15] H. Ullrich et al., "A 100ps 9K-gate ECL masterslice," in ISSCC Dig. Tech. Papers, pp. 200-201, 1985.

[16] T.H. Ning et al., "Self-aligned npn bipolar transistors," in IEDM Tech. Dig., pp. 823-824, 1980.
T.H. Ning et al., "Self-aligned bipolar transistors for high-performance and low-power-delay VLSI," IEEE Trans. Electron Devices, vol. ED-28, pp. 1010-1013, 1981.

[17] D.D. Tang et al., "1.25μm deep-groove-isolated self-aligned bipolar circuits," in ISSCC Dig. Tech. Papers, pp. 242-243, 1982.

[18] M. Suzuki et al., "A 165ps/gate 5000-gate ECL gate array," in Extended Abstracts 17th Conf. Solid-State Devices and Materials, pp. 377-380, 1985.

[19] S.K. Wiedmann and D.D. Tang, "High-speed split-emitter MTL/I²L memory cell," in ISSCC Dig. Tech. Papers, pp. 158-159, 1981.

[20] Y. Tamaki et al., "New U-groove isolation technology for high-speed bipolar memory," in Symp. VLSI Technology, Dig. Tech. Papers, pp. 24-25, 1983.

[21] K. Toyada et al., "A 15ns 16Kb ECL RAM with a pnp load cell," in ISSCC Dig. Tech. Papers, pp. 108-109, 1983.

[22] T.H. Ning and R.D. Isaac, "Effect of emitter contact on current gain of silicon bipolar devices," IEEE Trans. Electron Devices, vol. ED-27, pp. 2051-2055, 1980.

[23] J.M.C. Stork and J.D. Cressler, "The impact of non-ohmic polysilicon emitter resistance on bipolar transistor performance," in Symp. VLSI Technology, Dig. Tech. Papers, pp. 47-48, 1986.

[24] D.D. Tang and P.M. Solomon, "Bipolar transistor design for optimized power-delay logic circuits," IEEE J. Solid-State Circuits, vol. SC-14, pp. 679-684, 1979.

[25] P.M. Solomon and D.D. Tang, "Bipolar circuit scaling," in ISSCC Dig. Tech. Papers, pp. 86-87, 1979.

[26] T.H. Ning and D.D. Tang, "Bipolar trends," Proc. IEEE, vol. 74, pp. 1669-1677, 1986.

[27] J.M.C. Stork and R.D. Isaac, "Tunneling in base-emitter junction," IEEE Trans. Electron Devices, vol. ED-30, pp. 1527-1534, 1983.

[28] G.P. Li et al., "Identification and implication of a perimeter tunneling current component in advanced self-aligned bipolar transistors," to be published.

[29] S.A. Petersen and G.P. Li, "Hot-carrier effects in advanced self-aligned bipolar transistors," in IEDM Tech. Dig., pp. 22-25, 1985.

[30] T.H. Ning et al., "Scaling properties of bipolar devices," in IEDM Tech. Dig., pp. 61-64, 1980.

[31] D.D. Tang, "Heavy doping effects in pnp bipolar transistors," IEEE Trans. Electron Devices, vol. ED-27, pp. 563-570, 1980.

[32] T.C. Chen et al., "Reliability analysis of self-aligned bipolar transistors under forward active current stress," in IEDM Tech. Dig., pp. 650-653, 1986.

[33] D.D. Tang et al., "Design considerations of high-performance narrow-emitter bipolar transistors," IEEE Electron Device Lett., vol EDL-8, pp. 174-175, 1987.

[35] D.D. Tang et al., "On the impurity profiles of down-scaled bipol... transistors," in IEDM Tech. Dig., pp. 412-415, 1986.

[35] Y. Tamaki et al., "A 100nm emitter transistor fabricated with direct EB writing for high-speed bipolar LSI," in Symp. VLSI Technology, Dig. Tech. Papers, pp. 31-32, 1987.

PABLO vs. Double-Poly - A COMPARISON OF TWO HIGH-PERFORMANCE BIPOLAR TECHNOLOGIES

R.A. van Es and D.J.W. Noorlag

Philips Research Laboratories
P.O. Box 80000, 5600 JA Eindhoven
The Netherlands

1. INTRODUCTION

In this paper a comparison is made between two high performance bipolar processes with self-aligned features.

2. TECHNOLOGY

2.1 Process steps of PABLO

In the PABLO process (figure 1) a stack of lpcvd layers is deposited. With nitride and oxide as a mask the second polysilicon layer is etched and the sidewalls are laterally oxidized. (figure 1a).
The nitride is etched. Boron for the external base is implanted. The lateral oxide is removed and again a thin oxide is grown. Nitride is etched. (figure 1b).

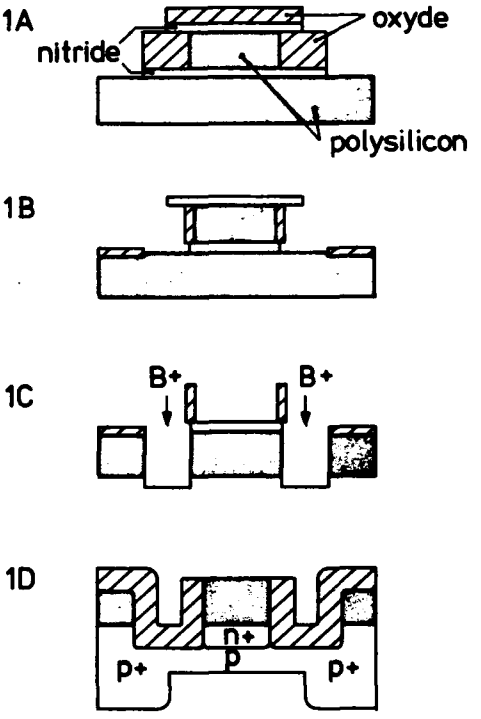


FIGURE 1
Main processing steps of the PABLO process



FIGURE 2
Cross sectional view of the PABLO transistor.

Next the exposed parts of the first and second polysilicon layer are etched. Boron is implanted into the groove (figure 1c).
The groove is oxidized and nitride is etched on the emitter area. Boron is implanted into the polysilicon. After base drive-in the emitter is implanted and driven in. (figure 1d)

2.2 Process steps of Double Poly

In the double poly process (figure 3), a polysilicon layer is deposited and heavily implanted with boron. Then an oxide layer is deposited and the polysilicon area's are defined. Both the oxide and polysilicon layers are etched by RIE. A thin oxide is grown and the intrinsic base is implanted through this oxide. (figure 3a).
Another oxide layer is deposited and oxide spacers are formed by anisotropic etching. (figure 3b).
A second polysilicon layer is deposited and arsenic is implanted. The second polysilicon layer is patterned (figure 3c) Then the emitter drive-in is performed.

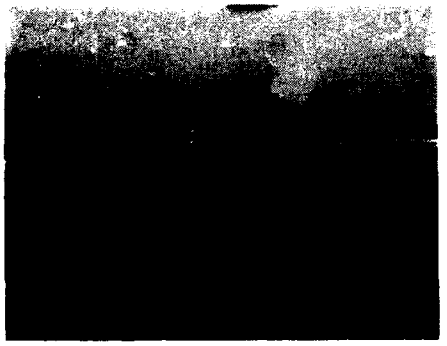


FIGURE 4

Cross sectional view of the Double Poly transistor

3. TRANSISTOR FEATURES

3.1 External Base

In both processes the external base is within submicron range of the intrinsic transistor. Moreover the emitter width is less then the smallest lithographic dimension.

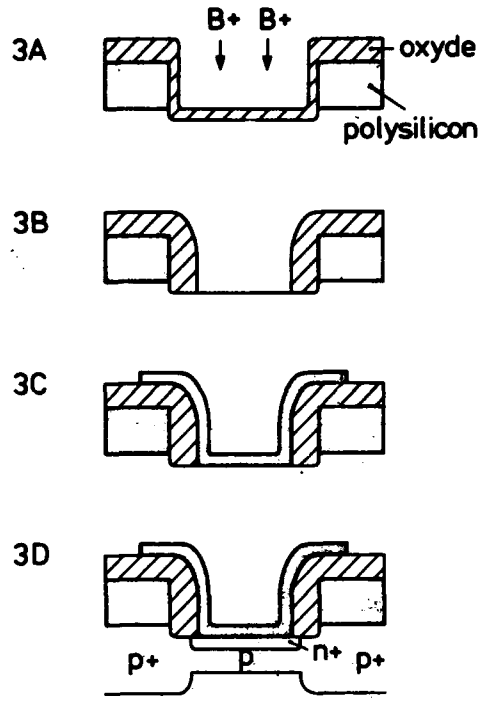


FIGURE 3

Main processing steps of the Double Poly process

3.2 Dry etching of Polysilicon

In the double poly concept the polysilicon is removed from the base area by a non-selective dry etching method. Due to the grain size of the polysilicon it could be difficult to obtain a smooth surface. In the PABLO process only the connection between intrinsic and extrinsic base is etched.

3.3 Emitter edge capacitance

When scaling down the emitter width the emitter-base edge capacitance becomes increasingly important. In the PABLO process the emitter edge is butted against the PABLO groove. Thus the emitter-base edge capacitance is significantly reduced.

3.4 Premature punchtrough

In order to prevent premature punchthrough it is necessary to avoid a local high base resistance rim around the emitter edge. (2) This means that a hookup implantation is necessary under the spacer, unless the spacer width is less then 0.2 to 0.3 micron. In the double poly concept this implantation is the same as the implantation for the intrinsic base. In the PABLO concept this implantation is independent of the intrinsic base area.

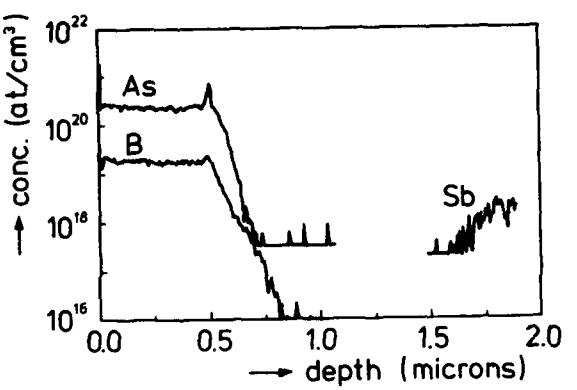


FIGURE 5

SIMS profile of a Double diffused transistor. The basewidth and emitterdepth are both approximately 0.15 micron.

3.5 Implantation Damage

In order to prevent implantation damage it is best to use polysilicon as a diffusion source. (double diffused) In this case in the double poly concept the link between intrinsic and extrinsic base must be provided by lateral diffusion. But making the hookup area to small leads to excess lateral diffusion of the baseboost. In the PABLO concept the dope of the hookup area can be tuned.

REFERENCES

1 S.F. Chu e.a. IBM General Technological Division, East Fishkill Hopewell junction, New York, Proc. 1e Int Symp. on VLSI Science and Technology '82 Electrochemical Soc.

2 D.D. Tang, G.P. Li, C.T. Chuang, T.H. Ning, "On the impurity profiles of Down Scaled Bipolar transistors" IBM T.J. Watson Research Centre, Yorktown Heights, N.Y. 10598

3 H.G.R. Maas and J.A. Appels, "PABLO, a versatile VLSI technology", Philips j. Res. vol. 39 no. 3 pp. 103-108, 1984

RAPID ANNEALING FOR SHALLOW JUNCTION FORMATION

A. E. Michel

IBM Thomas J. Watson Research Center
Yorktown Heights, NY 10598, USA

The application of rapid thermal annealing (RTA) of ion implanted impurities for the formation of shallow junctions provides advantages of reducing dopant distribution while removing the lattice damage. The higher temperature processing also improves the electrical activity of the dopants. The physical mechanisms of the redistribution and activation of ion implanted impurities are complex and RTA provides a valuable tool for investigating the details.

The scaling of silicon devices to smaller dimensions requires new methods for fabricating the necessary shallow structures. Bipolar transistors in particular require in addition very precise dopant control which is provided only by the ion implantation process. The removal of the damage to the crystalline lattice produced by the ion implantation process as well as the activation of the dopant require thermal annealing during which a diffusive redistribution of the dopant atoms occurs. While diffusion may be reduced either by reducing the temperature or reducing the time of the anneal, the two methods are not equivalent. For example, it was pointed out by Sedgwick [1] and subsequently demonstrated experimentally by Seidel [2] for arsenic implants, that the removal of extended defects involves the transport of silicon atoms which requires a higher activation energy (~5 ev) than the diffusion of of the arsenic impurities (~4 ev). Defect removal for the case of boron implantation is also governed by a 5 ev activation energy [3], while the boron diffusivity is activated by only 3.5 ev. Thus to minimize the dopant motion and still remove the implantation damage demands annealing at high temperature for short times.

Perhaps a more important factor that favors the high temperature short time process for the case of high impurity concentration is that the solubility limit increases with temperature. This applies to all the common dopants for silicon. Figure 1 shows the solubility limit plotted vs reciprocal temperature for arsenic and boron. The important characteristic for device performance is the electrical solubility which may be much less than the physical solubility. The "clustering" phenomenon of arsenic in silicon limits the electrical activity to concentrations well below the physical solubility limit at a given temperature. An increase by a factor of 3 in electrically active arsenic may be achieved by annealing at 1150°C rather than 800°C. Wilson [4] has reported that high temperature rapid anneals produce similarly large improvements in the conductivity of arsenic doped polycrystalline silicon, which is particularly relevant for bipolar transistors with arsenic doped polysilicon emitters. As seen in Fig. 1 the electrical solubility of boron exhibits an even greater temperature dependence than that of arsenic. In addition, for ion implanted boron annealed at low temperature, a major fraction of the boron may be electrically inactive even though the maximum concentration is well below the equilibrium electrical solubility limits of Fig. 1 [5]. For this condition the boron is probably trapped on defects which have not been removed by the low temperature anneal. In any case the sheet resistance is greater for a low temperature anneal than one performed at higher temperature.

Another factor that must be considered is that of an anomalous displacement produced by the anneal of ion implanted samples. For arsenic the majority of the data is capable of being fit with accepted diffusion mech-

anisms and hence there does not appear to be an anomalous transient effect [6]. There are still some unresolved questions concerning the time constant for the clustering process which controls the amount of arsenic free to move at the beginning of the anneal cycle and hence on the amount of displacement obtained.

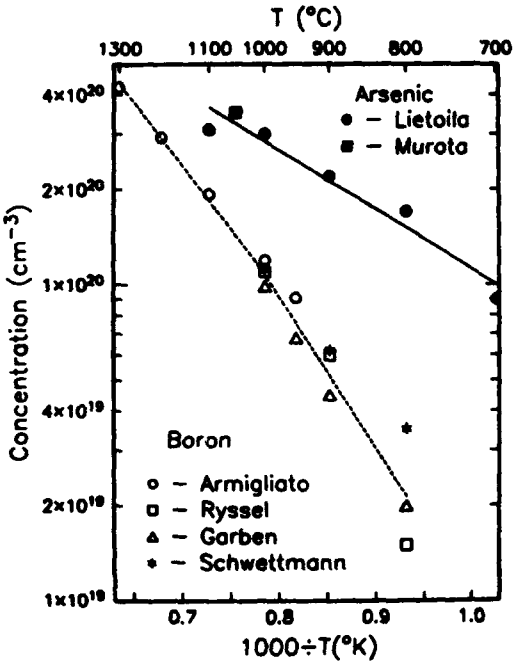


Figure 1

Impurity solubility vs 1/T.

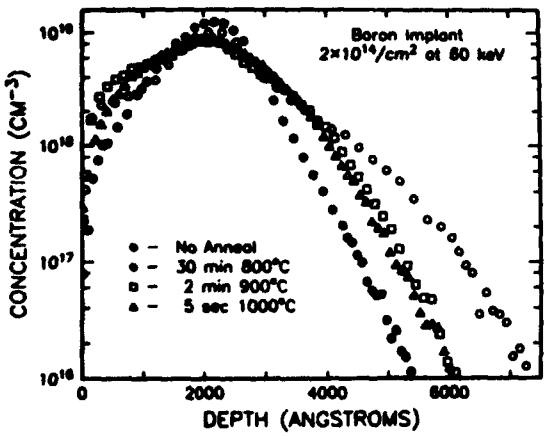


Figure 2

Boron transient diffusion at various temperatures.

Boron implants, on the other hand, may exhibit a large anomalous transient displacement depending on the anneal temperature [5,7-9]. Fig. 2 shows profiles of implanted boron before and after annealing at several temperatures between 800 and 1000°C. An anneal at 800°C for 35 minutes may produce a displacement in the low concentration tail region of the order of 200 nm. By annealing at higher temperatures, ~1000°C, for 5 seconds, the anomalous displacement of the junction is diminished to ~50 nm. Also, as mentioned above, the electrical activity of the layer is increased by the higher anneal temperature. The anomalous diffusion persists down to temperatures as low as 700°C as shown in Fig. 3. The shape of the depth profile also exhibits a dependence on the anneal temperatures. At 700 and 800°C relatively little motion of the boron occurs in the region of the peak and the profile forms bulges on both sides of the peak. As the anneal temperature is increased the bulges are displaced to higher concentrations, until at 1000°C all of the boron is involved in the redistribution. Thus as the temperature is increased a greater number of boron atoms participate in the anomalous diffusion. The electrically active concentration limit also coincides with the position of the bulges; it appears that the mobile boron is that which is electrically active [7]. The time duration of the anomalous transient also diminishes rapidly with increasing temperature from ~60 min at 800°C to ~1 sec at 1000°C. A complete understanding of the transient enhanced diffusion is still lacking, however, it is likely

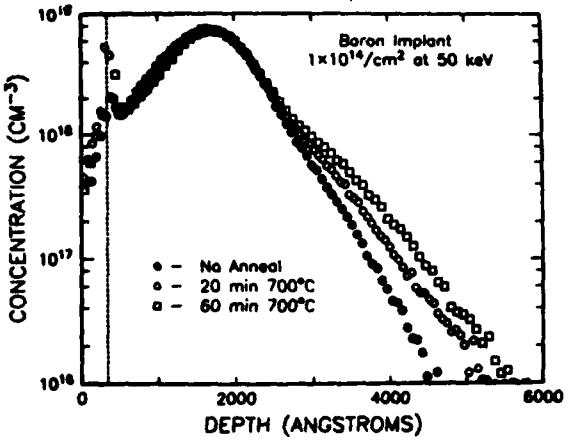


Figure 3

Transient diffusion at 700°C.

that the basic mechanism is caused by the dissolution of clusters of interstitial silicon atoms formed during the implantation process. Recent experiments on damage introduced by silicon implantation into stabilized boron profiles [10] support such a mechanism and also demonstrate another phenomenon, namely, the reduction of anomalous diffusion in the region of high damage. Fair [11] proposed that the formation and growth of extended defects sinks the silicon interstitial population which reduces the boron diffusion.

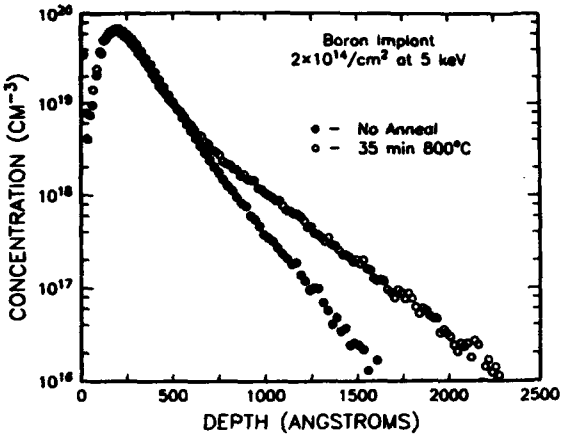


Figure 4

Boron channeling tail at low energy and transient diffusion.

The channeling along axial and planar crystallographic directions is a major contributor to the junction depth of boron layers formed by low energy ion implantation [12]. Fig. 4 shows the long channeling tail and the anomalous motion for a 5 keV boron implant. As demonstrated by Crowder [13], pre-amorphization of the crystalline lattice effectively eliminates the boron channeling tail. The existence of anomalous transient diffusion of boron implanted into amorphized silicon is controversial. Much of the controversy in regard to RTA is explained by a lack of knowledge of the sample temperature which is difficult to measure absolutely in a non-invasive way. Uncertainties of the thickness and abruptness of the amorphized layer, which for Si ion bombardment depends on the substrate temperature during implantation, also contribute to the disagreement. In a recent study by Sedgwick [14], in which special attention was given to temperature determination and amorphous layer thickness, the motion during RTA of boron implanted into amorphized silicon are fit by diffusion modeling without invoking anomalous transient effects. Fig. 5 shows a comparison between the measured and modeled profiles for two rapid annealed samples.

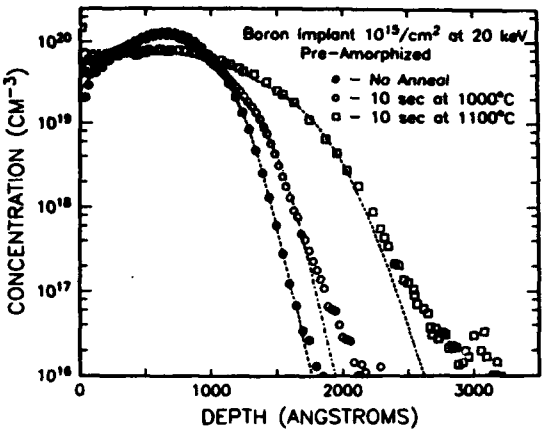


Figure 5

Boron diffusion in pre-amorphized Si.

The reported electrical activity of boron implanted into amorphized silicon also suffers from considerable disagreement. Comparing boron implanted into pre-amorphized and crystalline silicon, Yamada [15], using 1 hour furnace anneals, finds lower sheet resistance for pre-amorphized samples annealed at temperatures below 900°C and the same sheet resistance for samples annealed at higher temperatures. The resistances values indicate that during the solid phase epitaxial regrowth of the amorphous layer boron is incorporated substitutionally into the lattice at concentrations far in excess of the solubility limit. Sedgwick [14] finds boron precipitation and higher resistance in amorphized samples than in their crystaline counterparts for anneal temperatures above 950°. Vasudev [16] obtained much lower sheet resistance values with 2 minute anneals at 600°C than with longer 30 minute 800°C anneals and still lower values with values with a two step anneal consisting of 110 sec at 600°C followed by 10 sec at 1100°C. The electrical activity of boron in recrystalized amorphized layers is likely to be sensitive to other impurities which may act as nucleation sites for precipitates.

A problem with the pre-amorphization technique for the production of shallow junctions is the failure to remove the extended defects formed just below the amorphous layer. Such defects are a particular hazard for bipolar transistors in that they are a potential source for "pipes" and may represent a significant yield detraction. Sands [17] has demonstrated defect-free regrowth using Ge ions for amorphization and RTA at 1150°C. Ajmera and Rozgonyi [18] also employed Ge ions and RTA to eliminate both end of range defects as well as the residual damage produced by implantation through a sloping SiO_2 mask edge. The determination of "defect free" was based on transmission electron microscope observations which servey only a small volume of the sample, and the possible degredation of yield for bipolar devices has yet to be evaluated.

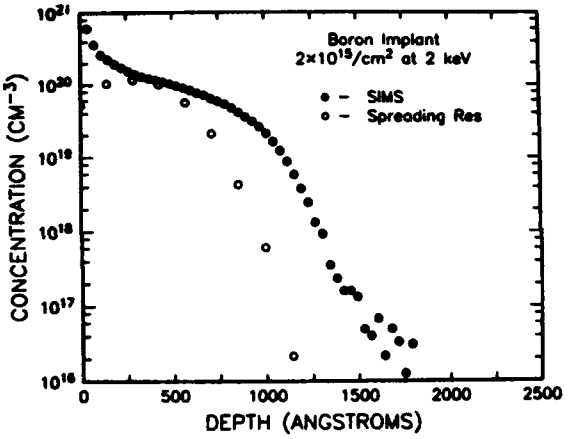


Figure 6

Chemical and carrier profiles for very low energy boron implants after 10 sec anneal at 1050°C.

Another technique reported by Davies [19] combining rapid thermal annealing with very low energy boron implants into crystalline silicon has shown considerable promise in the quest for very shallow junctions. Fig 6 shows a SIMS profile as well as a carrier concentration profile obtained by the spreading resistance technique. The junction depth determined from the SIMS profile is considerably greater than that obtained from the spreading resistance profile. It is generally observed that the junction depth obtained from carrier profiles, whether determined by spreading resistance or differential conductivity measurements or from simple bevel and stain measurements are generally shallower than the SIMS chemical profiles for boron. A satisfactory explanation for this is still lacking.

I wish to thank M. Demeo for performing the implantation, G. Scilla for providing the SIMS profile, and E. Gorey for providing the spreading resistance data on the 2 keV boron sample.

References

1. T.O. Sedgwick, J. Electrochem. Soc. 130 (1983) 484.
2. T.E. Seidel, D.J. Lischner, C.S. Pai, R.V. Knoell, D.M. Maher, and D.C. Jacobson, Nuclear Instruments and Methods in Physics Res. B7/8 (1985) 251.
3. D. Harame, S. Mader, A. Michel, and J. Cressler, (to be published)
4. S.R. Wilson, R.B.Gregory, W.M. Paulson, S.J. Krause, J.D. Gressett. A.H. Hamdi, F.D. McDaniel, and R.G. Downing J. Electrochem. Soc. 132 (1985) 922.
5. A.E. Michel, W. Rausch, P.A. Ronscheim, and R.H. Kastl, Appl. Phys Lett. 50 (1987) 416.
6. T.O. Sedgwick, A.E. Michel, S.A. Cohen, V.R. Deline, and G.S.Oehrlein , Appl. Phys. Lett. 47 (1985) 848.
7. W.K. Hofker, H.W. Werner, D.P. Oosthoek, and N.J. Koeman, Appl. Phys. 2 (1973) 265.
8. R.T. Hodgson, V. Deline, S.M. Mader, and J.C. Gelpey, Appl. Phys. Lett., 44 (1984) 589
9. M. Servidori, R. Angelucci, F. Cembali, P. Negrini, and S. Solmi, and P. Zaumseil and U. Winter, Jour. Appl. Phys. 61 (1986) 1834.
10. A.E. Michel, W. Rausch, and P.A. Ronsheim, Appl. Phys. Lett. (in press).
11. R.B. Fair, J.J. Wortman, and J. Liu,J. Electrochem. Soc. 131, (1984) 2387.
12. A.E. Michel, R.H. Kastl, S. Mader, B.J. Masters and J.A.Gardner, Appl. Phys. Lett. 44 (1984) 404.
13. B.L. Crowder, J.F. Ziegler, and G.W. Cole, in: Crowder, B.L. (ed), Ion Implantation in Semiconductors and Other Materials, (Plenum Press, New York, 1973) pp 257-265.
14. T.O. Sedgwick, A.E. Michel, V.R. Deline, S.A. Cohen, and J.B. Lasky, (to be published).
15. K. Yamada, M. Kashiwagi, and K. Taniguchi, J. Jour Appl. Phys 22 (1983) 157.
16. P.K. Vasudev, A.E. Schmitz, and G.L Olson, Mat. Res. Soc. Symp. Proc. 35 (1985) 367.
17. T. Sands, J. Washburn, E. Myers, and D.K. Sadana, Nuclear Inst. and Methods in Phys. B7/8 (1985) 337.
18. A.C. Ajmera and G.A. Rozgonyi, Appl. Phys. Lett. 49 (1986) 1269.
19. D. Eirug Davies, IEEE Elec. Dev. Lett. EDL-6 (1985) 397.

Session B2.1

SOI Workshop I

Chairman: D. Mc Caughan

Tuesday, September 15, 1987

RECRYSTALLIZATION OF SILICON-ON-INSULATOR FILMS BY A PSEUDOLINE ELECTRON BEAM

Hiroshi ISHIWARA and Susumu HORITA

Graduate School of Science and Engineering, Tokyo Institute of Technology
Nagatsuda, Midoriku, Yokohama 227, Japan

Recent progress on the recrystallization of silicon-on-insulator (SOI) films by a pseudoline electron beam is reviewed. The pseudoline beam was produced by scanning a spot beam along a line faster than the thermal response time of the substrate. In order to obtain a large SOI region without sub-boundaries and voids, such recrystallization conditions as the scanning waveform to produce a psuedoline beam, the scanning direciton and velocity of the pseudoline beam, the seed direction, and so on were optimized. A single crystal SOI area of 100 μm square was thus obtained.

1. INTRODUCTION

Electron beam (e-beam) recrystallization is one of the most promising techniques to fabricate silicon-on-insulator (SOI) structures suitable for three-dimensional integration. Particularly, use of a pseudoline e-beam [1-6], in which a spot beam is scanned along a line faster than the thermal response time of the substrate, is interesting from a viewpoint of productivity, as well as use of a line beam [7-9]. The scanning waveform to produce a pseudoline beam was initially a sinusoidal wave, which gave an increasing temperature profile towards both edges of the line and recrystallized the molten zone from its center as a single crystal [2,3]. More recently, a new waveform, an amplitude modulated sinusoidal wave, was proposed and successfully applied to recrystallization of large area SOI films [4].

However, there are still a few problems in the pseudoline e-beam recrystallization method, which are similar to those in the line beam method. The first one is generation of sub-boundaries which is inherently observed in the recrystallized large SOI region, and the second one is generation of voids in the SOI film, which is pronounced near the seed regions where the Si film is directly in contact with the Si substrate. In this paper, we review the recent progess of the studies on their generation mechanisms and suppression methods.

2. EXPERIMENTAL PROCEDURE

A schematic diagram of the electron beam annealing system used in this experiment is shown in Fig.1. A pseudoline electron beam abut 450 μm long was synthesized by scanning a spot beam with about 500-kHz sinusoidal or triangular signals which were applied to the X and Y electrostatic deflection plates. The acceleration voltage and the diameter of the spot beam were 10 kV and 240 μm, respectively. The pseudoline beam was then scanned electromagnetically as a whole with velocities from 0.1 to 10 cm/s. The oblique angle θ between the scanning direction of the pseudoline beam and the line normal direction was changed from 0° to 60°by changing the ratio of the X and Y amplitudes.

In the sample preparation, 1-μm-thick SiO_2 films were first deposited on Si(001) wafers using plasma-enhanced chemical vapor deposition (P-CVD). Then, the oxide films were partially etched to form stripe or square SOI regions with 50 to 200 μm in dimension. The shape of the seed region was either a continuous stripe 5 μm wide or a perforation seed structure, in which rectangular seed patterns are separately arranged along a line. The seed regions were aligned to <100>, <130>, or <110> axes of the substrate. On the patterned wafers, Si films 0.6 μm thick were vacuum deposited at temperatures around 500°C. Finally, the samples were capped with 0.5 μm-thick P-CVD oxide films and

mounted on a carbon holder kept at 500°C using Pb-Bi (80:20 wt%) alloy.

After recrystallization, the cap oxide was etched and the samples were dipped in Wright etchant to make the sub-boundaries in the Si films clear. The surface morphology of the samples was observed with Nomarski optical microscope. The samples were further characterized using the electron channeling contrast in SEM (scanning electron microscopy).

3. LATERALLY SEEDED EPITAXY

In the experiment of laterally seeded epitaxy of SOI films, the pseudoline beam was mainly scanned along a seed stripe, while the direction of the seed stripe to the Si substrate and the oblique angle θ were changed. Figure 2 shows Nomarski optical micrographs of Wright etched SOI films which were recrystallized by a pseudoline beam with a velocity of 10cm/s [5]. In this figure, comparison of (a) and (b) gives a variation of the sub-boundary-free area with the oblique angle θ in the samples having seed stripes along the <110> direction, while comparison among (b), (c), and (d) gives a variation with the seed direction at a fixed oblique angle of 45°.

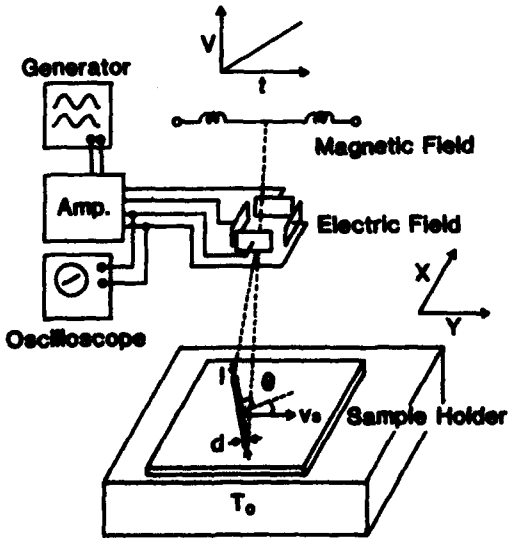


FIGURE 1

Schematic diagram of an e-beam annealing system

We can see from these figures that a laterally grown area with no sub-boundary is larger in the right-hand side of each SOI stripe than that in the left-hand side, since the lateral growth toward the left-hand side is enhanced due to the movement of the oblique S-L (solid-liquid) interface. We can also see from comparison of (a) and (b) that the sub-boundary-free growth length L_{sf} from the right-hand side is increased with the oblique angle θ, and from (b), (c), and (d) that the L_{sf} is increased in the sequence of the <110>, <130>, and <100> directions of the seed stripes. In the center stripe in Fig.2(d), the sub-boundary-free area covers the whole SOI region forming a grain boundary parallel to the seed stripe near the

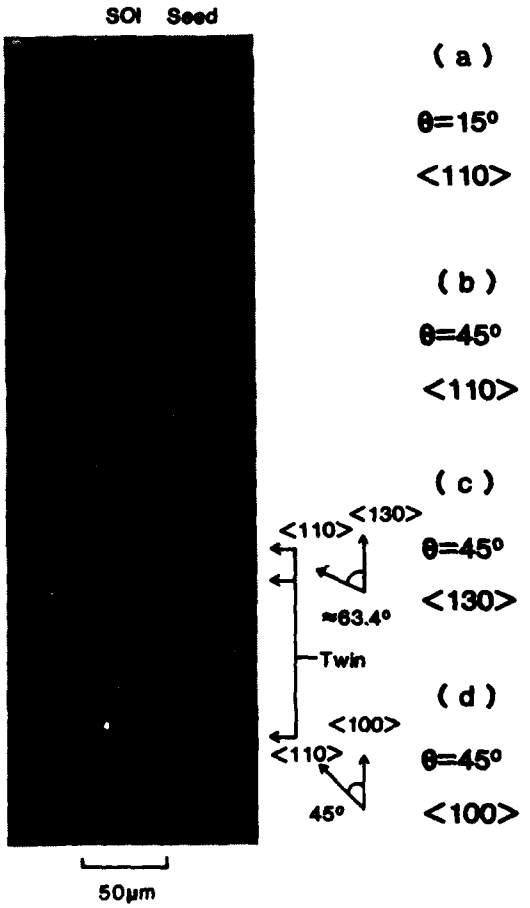


FIGURE 2

Nomarski optical micrographs of Wright etched sample. The seed direction and oblique angle are shown in the right-hand side in each figure.

left-hand edge, although microtwins along the <110> axis exist in the right-hand region.

The relation between the oblique angle θ and the sub-boundary-free growth length L_{sf} for the three samples with different seed directions is summarized in Fig.3 [5]. We can confirm from this figure that the lateral growth length is enhanced either by increasing the oblique angle θ or by rotating the seed direction from <110> to <100> axes. It was also found that L_{sf} was increased with increase of the beam power.

Next, properties of sub-boundaries are characterized. As can be seen from Figs.2(a) and (b), in the samples with <110> seed stripes the direction of sub-boundaries is roughly fixed at 45° from the seed stripes, even if the oblique angle θ is changed from 15 to 45°. The direction of sub-boundaries hardly changed in any samples with the <110> seed stripe. On the other hand, we can see from Figs.2(b), (c), and (d) that the direction of sub-boundaries is changed by the crystal orientation of the seed stripes, even if θ is kept constant. From these and other experimental results, we speculated that the direction of sub-boundaries is roughly aligned to the <100> direction of the Si substrate, because of the folded (111) facets formed at the solidification front [10].

In order to confirm the above speculation, SOI islands of 100 μm square which were surrounded by 5 μm-wide seed stripes were recrystallized by a nonoblique pseudoline e-beam. That is, a pseudoline beam was set parallel to an edge of the square pattern and it was scanned to the normal direction of the line.

Figure 4 shows electron channeling contrast micrographs in the upper region of SOI islands which were recrystallized by scanning a pseudoline beam upwards [5]. The top portion in the micrograph is the seed stripe and because of its heat sink effect the edge region of the SOI island was solidified downwards. We can say that the lateral growth length from the top seed stripe is not so long as to form (111) facets, while the growth length from the bottom seed stripe (about 100 μm) is long enough to form (111) facets at the S-L interface. Thus, the shape of the grain boundary which is formed by collision of the two solidification fronts

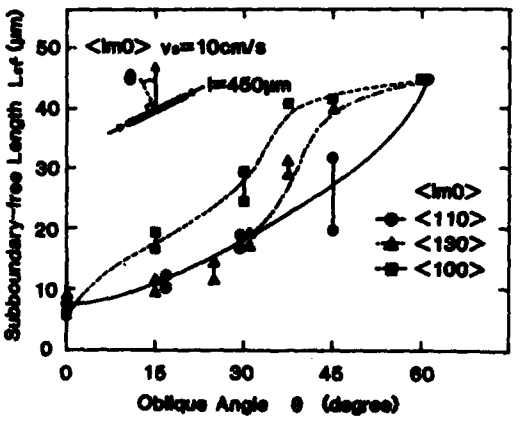


FIGURE 3

Variation of the sub-boundary-free growth length L_{sf} with the oblique angle θ for three seed directions.

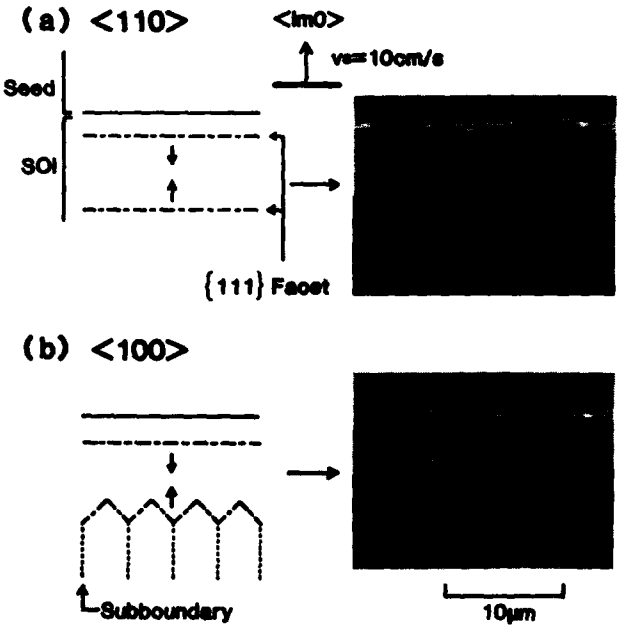


FIGURE 4

Electron channeling contrast micrographs of Wright etched samples recrystallized by a non-obliquely scanned e-beam, and schematic drawings of the S-L interface during solidification.

is considered to reflect the shape of the (111) facets at the S-L interface which goes upwards.

We can see from this figure that the grain boundary is flat when the beam is scanned along the <110> direction, while it exhibits a round saw-tooth shape when it is scanned along <100>. These shapes are consistent with the shapes of (111) facets in respective seed directions as shown in the left-hand side. We conclude from this experiment that folded (111) facets are generated at the S-L interface when the beam is scanned along the <100> direction, and sub-boundaries are generated at the interior corners of the folded (111) facets.

4. PERFORATION SEED STRUCTURE

It has been reported that voids or holes are often generated in SOI films which are recrystallized by line or pseudoline e-beams when the beam crosses seed stripes where the Si film is directly in contact with the Si substrate[1,8]. Under certain conditions, the voids are as large as several hundred microns, and they often produce new defects such as grain boundaries in the film. Thus, it is impossible to fabricate integrated circuits in these films, unless the void generation is suppressed.

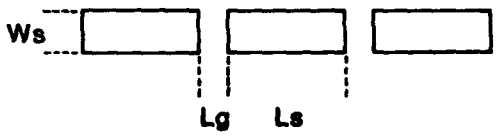


FIGURE 5

Schematic diagram of perforation seed structure

We studied the generation conditions of such voids using a pseudoline electron beam and speculated that the voids are generated by Si or SiO gas which is vaporized by abrupt, local temperature change and contained between the capping and underlying SiO_2 films, where SiO gas is known to be brought by dissolution of oxide films. Based on this speculation, we proposed a perforation seed structure [6], in which rectangular seed regions are separately arranged along a line, so that the heat sink effect is not so different between the seed and SOI regions and voids are not generated.

The size of seed regions ranged from 5 to 20 μm in width (W_s) and 5 to 100 μm in length (L_s) as shown in Fig.5. The seed regions were aligned along a line with a constant gap length L_g between the adjacent seed regions. L_g was changed from 3 to 15 μm in different samples.

Figures 6(a) and (b) show the electron channeling contrast micrographs of recrystallized SOI films with a conventional seed stripe and with a perforation seed structure, respectively [6]. The sizes of the structure are $L_g=W_s=5\,\mu m$ and $L_s=10\,\mu m$. A pseudoline e-beam was scanned

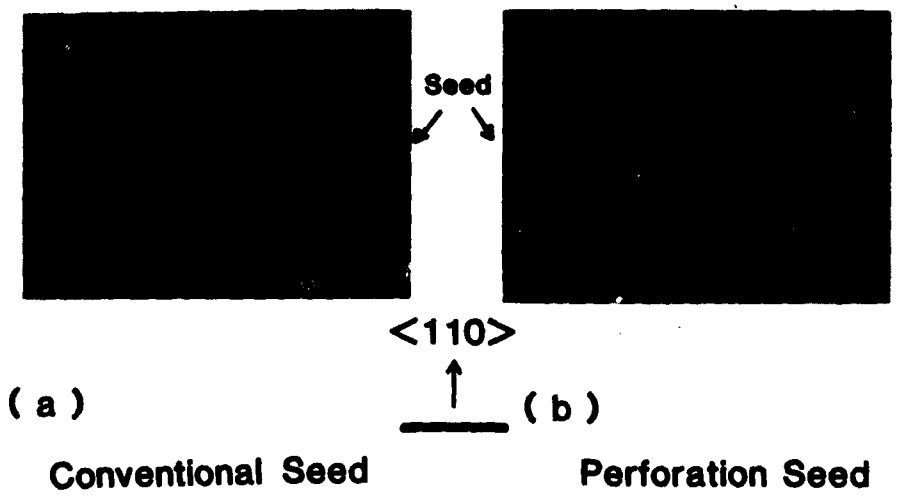


FIGURE 6 Electron channeling contrast patterns for SOI samples with (a) conventional and (b) perforation seed structures

perpendicular to the seed line from bottom to top of the figures. We can see from Fig.6(a) that two large voids as well as a small one are generated in the upper SOI region near the seed stripe. On the contrary, in the case of the perforation seed structure in Fig.6(b), no void is generated in the SOI film. We can also see that though the seed regions are separately arranged, their heat sink effect is strong enough to stop propagation of the randomly nucleated Si grains across the seed line. Thus, the crystalline quality of the upper SOI film is kept excellent up to about 30 μm from the seed line by laterally seeded epitaxial growth. However, the channeling contrast pattern shows that fluctuation of the crystal orientations which reflects the seed structure occurs beyond that region.

The size of the perforation seed structure was optimized from several experiments. It was found that a shorter L_g gives a better result from a viewpoint of the crystalline quality of the SOI film, however the minimum value for W_s= 5 μm is about 5 μm. Concerning the seed length L_s, a longer value was found to be more effective for controling the crystal orientations of the SOI film. Thus, the optimum value of L_s is considered to be 30 to 50 μm. These values will be somewhat modified if the seed width W_s is changed. For example, if W_s is narrower than 5 μm, the heat sink effect of the seed line becomes weaker and we will be able to choose the shorter L_g and the longer L_s without void generation. The seed regions with W_s=1 μm are expected to be still effective for stopping propagation of randomly nucleated Si grains.

5. LARGE AREA RECRYSTALLIZATION

We can conclude from the previous discussions that, in order to recrystallize a large SOI area in a single crystal, it is necessary to suppress the void generation as well as to suppress the (111) facet formation at the S-L interface. The former problem was found to be solved by use of the perforation seed structure. Concerning the facet formation, it was found that the (111) facets are likely to be suppressed when the S-L interface is aligned to the <$\bar{1}$10> direction and the beam is scanned along the <110> direction, as shown in Fig.4(a). Thus, the perforation seed structure was formed along the <$\bar{1}$10> direction and in some samples <110> seed stripes were also formed in order to stop the penetration of grain boundaries from the beam edge region.

Furthermore, in order to keep the S-L

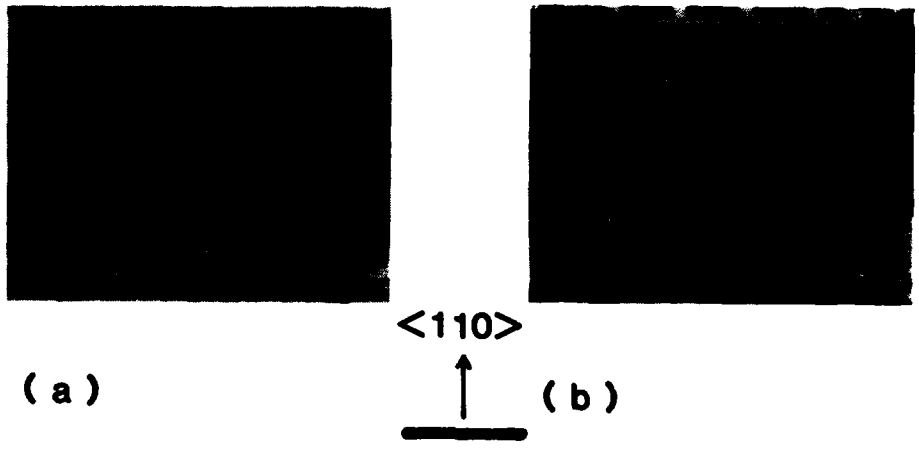


FIGURE 7 Electron channeling contrast patterns for SOI samples recrystallized by (a) conventional and improved methods

interface at the trailing edge of the beam as straight as possible, a pseudoline e-beam was synthesized by scanning a spot beam with a combined waveform of triangular and sinusoidal waves. This method was conceived from a fact that the molten zone shape becomes convex at the trailing edge when a spot beam is scanned with a triangular wave, while it becomes concave when the beam is scanned sinusoidally. The optimum amplitude ratio of the triangular and sinusoidal waves was determined to be 1 to 4 from the recrystallization experiment.

Figure 7 shows a comparison of the electron channeling contrast patterns for 100 μm square SOI regions recrystallized by conventional and improved methods. That is, in the sample shown in Fig.7(b), the perforation seed structure and the scanning using a combined waveform were employed, instead of the seed structure of a continuous stripe and the sinusoidal scanning in Fig.7(a). It is clear from the figures that no void is generated in the SOI region in Fig.7 (b), though generation of voids and crystalline defects is pronounced in Fig.7(a). The uniform contrast pattern in (b) also shows that the crystal orientation is well controlled except a small region of the upper left corner.

6. SUMMARY

We investigated the generation mechanisms of sub-boundaries and voids in the pseudoline electron beam recrystallization to propose a novel recrystallization method and a sample structure. Main results obtained are as follows.

1) In laterally seeded epitaxy of an SOI region from a seed stripe, the sub-boundary-free area is enhance either by increasing the oblique angle θ of the pseudoline beam, or by rotating the seed directions from <110> to <100> axes of the substrate.
2) Folded (111) facets are generated at the S-L interface when the pseudoline e-beam is scanned along the <100> direction, and sub-boundaries are generated at the interior corners of the folded facets.
3) A perforation seed structure in which rectangular seed regions are separately placed along a line is effective for stopping propagation of randomly nucleated Si grains as well as suppressing the void generation.
4) An SOI region of 100 μm square was recrystallized in a single crystal after optimization of the seed structure, the scanning direction, and the scanning waveform.

ACKNOWLEDGMENT

The authors gratefully acknowledge useful discussion with Professor S.Furukawa. They are also thankful to the staffs in JEOL Ltd. for their technical supports.

REFERENCES

[1] Davis, J.R., McMahon, R.A., and Ahmed, H., J. Electrochem. Soc. 132 (1985) 1919

[2] Ishiwara, H., Nakano, M., Yamamoto, H., and Furukawa, S., Proc. of 14th Conf. on Solid State Devices, Tokyo, 1982 [Jpn. J. Appl. Phys. Suppl. 22-1 (1983) 607]

[3] Ishiwara, H., Ohyu, K., Horita, S., and Furukawa, S., Jpn. J. Appl. Phys. 24 (1985) 126

[4] Hamasaki, T., Inoue, T., Higashinakagawa, I., Yoshii, T., and Tango, H., J. Appl. Phys. 59 (1986) 2971

[5] Horita, S. and Ishiwara, H., J. Appl. Phys. 61 (1987) 1006

[6] Horita, S. and Ishiwara, H., Appl. Phys. Lett. 50 (1987) 748

[7] Hayafuji, Y., Yanada, T., Usui, S., Kawado, S., Shibata, A., Watanabe, N., and Kikuchi, M., Appl. Phys. Lett. 43 (1983) 473

[8] Knapp, J.A., J. Appl. Phys. 58 (1985) 2584

[9] Hada, H., Saitoh, S., and Okabayashi, H., Extended Abstract of 17th Conf. on Solid State Devices and Materials, 1985 (Business Center for Academic Societies, Tokyo, 1985) p.139

[10] Geis, M.W., Smith, H.I., Tsaur, B-Y., Fan, J.C.C., Silversmith, D.J., and Mountain, R.W., J. Electrochem. Soc. 129 (1982) 2812

A COMPARATIVE STUDY OF STARTING MATERIALS FOR SOI DEVICE FABRICATION OBTAINED BY DIFFERENT RECRYSTALLIZATION PROCEDURES AND MATERIAL STRUCTURES

D.J.WOUTERS, M.R.TACK, P.V.MERTENS, H.E.MAES and C.L.CLAEYS

IMEC vzw, Kapeldreef 75, B-3030 Leuven, Belgium

Different conditions of zone melting recrystallization using both laser and a mercury arc lamp are studied. N-channel MOSFET's were fabricated in these materials to compare their qualities and to evaluate their device-worthiness. Subgrainboundary-free silicon films are obtained resulting in devices with high surface mobility and low leakage currents.

1. MATERIAL PREPARATION

100 mm (100) oriented CZ silicon wafers were prepared for recrystallization experiments. Seeding areas parallel to the [110] direction with 0.5 µm thick oxide islands were defined by a LOCOS technique.

For the laser recrystallization a 0.5 µm LPCVD polycrystalline silicon film was deposited. Four different capping layers were used in the experiments : 6 nm or 55 nm silicon nitride, 35nm oxynitride, and a periodic structure of 55 nm thick and 7 µm wide (anti-reflective) silicon nitride stripes with 10 nm silicon nitride in between. The spacing between the stripes was 18 µm. The 55 nm silicon nitride films were annealed for 30 min in N_2 at 1100 °C

For the lamp recrystallization a 5 µm thick LPCVD poly-silicon was deposited. A 3 µm PECVD silicon dioxide layer was used as the capping layer.

2. RECRYSTALLIZATION CONDITIONS

A CW Argon ion laser was used for the laser recrystallization. The scanning of the circular laser spot was performed with galvanometer driven mirrors. A unidirectional scan method together with a large overlap between successive scan lines was applied resulting in a semi-continuous crystal growth [1].

For each one of the used combinations, optimal scanning parameters were determined experimentally (Table I).

The lamp recrystallization [1] was done with a capillary mercury lamp which in these experiments operates at an electrical power of around 350 W/cm. The complete topheater unit was scanned at 0.4 mm/s along the <110> direction over a wafer preheated to about 1275 °C.

3. QUALITY OF THE RECRYSTALLIZED SILICON

3.1. Laser recrystallization

For the laser recrystallized material mass transport occurs at the seeding edges of the

TABLE I

LASER RECRYSTALLIZATION CONDITIONS AND SURFACE MOBILITIES

CAP LAYER	LASER POWER [W]	SCAN VELOCITY [cm/s]	VERTICAL STEP [μm]	SURFACE MOBILITY [cm^2/Vs]
6nm Si_3N_4	17	10	15	390
55nm Si_3N_4	5.5	10	17	300
Stripes	12.5	50	5	550
35nm SiN_xO_y	7	10	15	390

Laser spot size is 100μm ($1/e^2$ intensity points)
Wafer preheating temperature is 350°C

oxide islands. The amount is found to depend on the type of capping layer. For the thin 6 nm nitride cap layer considerable mass transport is observed, eventually resulting in voids at the "entry" corner of the oxide islands, whereas with the thick 55 nm nitride and the 35 nm oxynitride capping layer the effects are much less pronounced. It can be concluded that mass transport decreases if a more rigid capping layer is used and that it can be effectively reduced by using a stripe structure.

For the continuous capping layers, a (Sub)-GrainBoundary (SGB) structure is formed with grains parallel to each other and to the <100> direction, for a scan direction parallel to the [110] direction (figure 1). The distance between these SGB's varies between 1 to 10 μm, depending on the type of capping layer. For the periodical striped capping structure SGB free crystals are obtained extending a few 100 μm from the seeding edge (figure 2).

3.2. Lamp recrystallization

Due to the combination of a relatively large amount of seeding area and a relatively thin burried oxide the power window for the top-heater was very small. This resulted in some material transport.

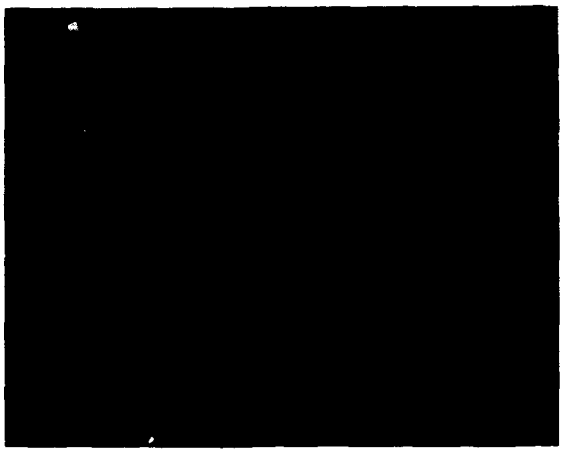


FIGURE 1

Photomicrograph of laser recrystallized material with continuous 55 nm nitride capping layer showing typical grainboundary structure. The silicon film was Secco etched. Distance between the grainboundaries is of the order of 1 μm. Grainboundary-free seeded distance is about 10 μm. The length of the marker is 50 μm. Scanning was from right to left and from bottom to top.



FIGURE 2

Photomicrograph of laser recrystallized material with anti-reflective stripes. Long (sub)grainboundary-free crystals are obtained. The length of the marker is 25 µm. Grain growth is from bottom (seed area) to top.

The major defect structure for the obtained silicon layers are (111) stacking faults.

The residual doping level of the obtained films ranges from $6*10^{14}$ to 10^{15} at./cm^3

4. DEVICE FABRICATION

Devices were processed in the recrystallized wafers, as well as in reference bulk-wafers using a 5 µm NMOS process, modified for thin SOI films.

The major characteristics of this process are: LOCOS-isolation, a 60 nm gate-oxide, Vt-adjust implant, B-implantations to suppress back-interface- and channel edge-leakage and As-P graded junctions. Finally dopants were activated by short time annealing.

5. ELECTRICAL CHARACTERIZATION

The electrical evaluation was performed on conventional and on five-terminal MOS-transistors.

5.1. Laser recrystallization

The threshold voltage of the front transistors is adjusted to about 1.3 V by the boron Vt implant. The surface mobility of both front- and back-transistors as a function of the different materials is listed in Table I. High front-channel mobilities in the order of 550 cm^2/Vs are obtained in the material with anti-reflective nitride stripes. Those values are comparable to the ones obtained in bulk silicon. The wafers that received a 6 nm nitride or a 35 nm oxynitride capping show somewhat lower mobilities and larger spreads whereas the worst values are obtained on the 55 nm nitride capped wafers. Those results correlate with the occurrence of SGB's.

Figure 3 shows the Ids-Vgs1 (front gate voltage) characteristic for a device processed in material with a striped capping structure. The diode leakage currents are less than 1 pA/(µm gate width) at a reverse bias of -10V. Interface state densities are around $7*10^{10}$ /cm^{2}eV and $6*10^{11}$ /cm^{2}eV for the front and back interface respectively, as measured by charge pumping. Figure 3 also shows the influence of the back-gate voltage (Vgs2) on Vt1 which indicates that the device operates partly in the thin-film regime. As a consequence, by applying an adequate Vgs2, the kink effect can be suppressed.

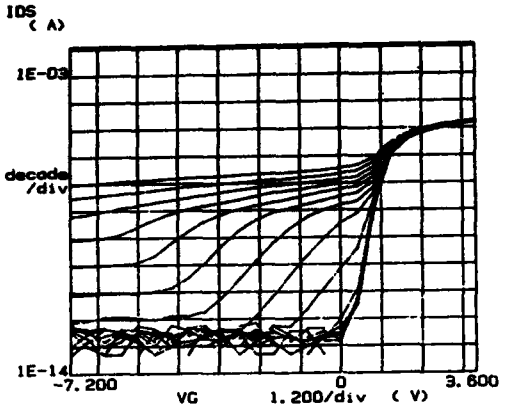


FIGURE 3

Log(Ids)-Vgs1 curves with Vgs2 as a parameter of a five-terminal thin film SOI MOSFET in laser recrystallized material with anti-reflective stripes. Device dimensions are W/L=24 µm/12 µm. Back gate voltage (Vgs2) varies from 9 V (upper curve) to -10 V with -1 V steps. Vds is 100 mV, and the film contact is grounded.

5.2. Lamp recrystallization

Surface roughness due to the applied polysilicon deposition technique is too high and results in poor gate oxide performance. Also local PECVD oxide cap failures resulted in pits. Nevertheless surface mobilities of around 500cm^2/Vs are measured. Reverse biased (-10V) diodes of 24 µm by 18 µm result in a leakage current below the detection threshold of 10^{-13} A.

6. CONCLUSIONS

Four different seeded material structures recrystallized by means of laser were evaluated. High surface mobility (550 cm^2/Vs) and low leakage currents are measured in material with an anti-reflective periodical capping layer, in which grainboundary-free material was obtained.

Lamp recrystallized thick film material shows comparable electrical results.

ACKNOWLEDGEMENT

Part of this work was performed within the ESPRIT project 370.

REFERENCES

[1] D.Wouters, P.Mertens, H.E.Maes in: The C.E.C.(ed.), ESPRIT 85 Status Report, part 1, (North Holland, Amsterdam, 1986) 203

3D-SOI INTELLIGENT POWER STRUCTURES

B. Dunne, C. G. Cahill, A. Mathewson, W. A. Lane

National Microelectronics Research Centre,
University College Cork,
Ireland.

M. Montier, D. Chapuis

Thompson EFCIS, 17 Avenue des Martyrs,
Grenoble, France.

1. INTRODUCTION

One area of the microelectronics industry which has attracted considerable research interest for a number of years is the concept of vertically stacking several planes of active devices for Three-Deimensional Integrated Circuits. While a number of means of achieving this have been suggested, the most common approach is to use stacked layers of energy beam recrystallised silicon, separated from each other and the bulk material by silicon dioxide. While an initial impression would suggest that the major benefits of such an approach would lie in packing density and speed performance it has been shown that the gains obtained in the context of a VLSI requirement are less than could be expected [1] when the extra complexity involved in producing the devices is considered [2,3].

However, the stacked SOI approach does offer significant advantages for the realisation of integrated, mixed technology systems, where separate layers of transistors can be individually optimised and separated by a high quality isolating oxide. It is clear that several varieties of radically different devices could be used on different levels of a 3D-SOI system, without seriously compromising the fabrication sequence of any of them. The layers, in principle, can be built up sequentially with only the thermal load of subsequent processing steps effecting previous stages.

Two alternative configurations of such an approach, fully stacked and mezzanine structures, are shown in Figure 1. The layered nature is clearly shown, with the vertically stacked structure being seen as a desirable longer term goal toward which the mezzanine approach used here is a logical stepping stone. In addition, a polysilicon shield or ground plane level could also be integrated within the vertically stacked structure and is included as a diffused region in the bulk silicon layer for the mezzanine approach, thus providing improved DC and transient isolation for the control logic array. For the vertically stacked approach this ground plane has the added benefit of being a useful heat sink during top layer recrystallisation and a means of achieving greater planarisation.

2. Intelligent Power/Interface IC

In evaluating this concept and its realisation, Intelligent Power/Interface applications were defined as being particularly suited to the 3D-SOI structure. Specifically, the high quality dielectric isolation offered by the inter-layer oxide offers great potential in this application, in addition to the capability of building optimized bulk silicon power devices and high performance latch-up free CMOS SOI circuits.

In order to provide the greatest design flexibility in a demonstration of a 3D-SOI intelligent power structure, gate array

implementations were adopted for both bulk power devices and the SOI control logic. The technologies chosen for this work are a 50V lateral DMOS bulk technology together with a 3um CMOS SOI process. These two processes are being combined to fabricate a small test-bed 3D-SOI gate array, suitable for semi-custom interfacing and medium current/voltage (1A/50V) driving applications. Packaging constraints currently limit total power dissipation in a full sized version of this array to the 5 watt level.

A 3um SOI CMOS process was developed, in e-beam recrystallised material, the characteristics of which are shown in Figure 2 [8]. This is typical of SOI CMOS processes described elsewhere [9], and consequently most attention is given to the more novel LDMOS technology in this paper.

3. Mezzanine IC Design and Processing

Recrystallisation techniques which are particularly suited to the 3D-SOI approach are laser [4] and electron beam [5,6] zone-melting-recrystallisation (ZMR) [2], because of their localised, rapid heating of the silicon. In their simplest form these techniques produce large grain polysilicon with randomly oriented crystallites and grain boundaries, leading to a large scatter in the characteristics of MOS devices fabricated in the material. In this work both techniques are being used, with a 'seeded' approach adopted to provide bulk quality silicon in the SOI layer. The seed structure used in the test-bed design has a 43um pitch; however, some encouraging progress is being made at both the laser and e-beam sites within this project in extending this distance to 60-80um. An improvement in this will clearly lead to greater packing densities and larger scope in the geometry and breakdown voltage of the power devices employed. Experiments have shown that 200V DMOS devices could be fabricated with a seed window spacing of 60um. Selective Epitaxial Growth (SEG) [1] is used to planarise the seed structure and to reduce the mass transport that is currently observed as a feature of the ZMR approach.

The CMOS SOI portion of the array is designed so that the gate modules and routing channels are located entirely between seed windows; this makes reasonably efficient use of the highly regular structure dictated by the seeding requirements. The design of these modules is more or less traditional; an example of a cell is shown in Figure 3. It includes a device of each type (P and N), as well as two through cell routing vias. Both metal and contact levels will be programmable and a single level metal scheme is used. In this example the ultimate packing density has been compromised somewhat to accommodate the project requirement of both laser and e-beam ZMR.

A major consideration in the case of the bulk technology is that the devices must fit completely between seed windows to avoid the melting into the bulk material that occurs during the ZMR step. This constraint on the size of the devices requires that a medium voltage [50-70V] DMOS technology be employed. This has been designed by modifying the traditional circular geometry of the LDMOS transistor to provide a device of the requisite dimensions. These transistors have been produced in an n-well CMOS process which was modified to produce the necessary Vt and breakdown voltage characteristics.

Experiments have been performed on the LDMOS devices to establish the effects of the recrystallisation step and subsequent SOI processing on the device characteristics. Initial results indicate that, provided the ZMR step is constrained to provide good quality seeded material without melting the bulk silicon under the isolating oxide [i.e. is within the power window], no detrimental effects are observed on bulk device performance [see table 1]. However, if the power window is exceeded some melting of the gate poly-silicon can be observed. Figure 4 shows a polysilicon track which exhibits mass transport effects due to high power ZMR processing.

4. Conclusions

A 50-70V bulk LDMOS process compatible with electron beam and laser ZMR for 3D-SOI application has been developed. It has been established that the devices produced do not degrade significantly in performance if care is taken to stay within the thermal budget and optimized ZMR cycles are used. This LDMOS technology is combined with a 3 micron SOI-CMOS technology in a mezzanine configuration, thus incorporating the merits of a 3D approach in a mixed technology system. The design and processing of this demonstration structure has been constrained to be similar to that expected to be used in a fully stacked device. In this way some of the technological challenges of the fully stacked structure have been dealt with in this work.

ACKNOWLEDGEMENTS

This work is partially funded by the ESPRIT 245 programme and we would like to thank our colleagues in Thomson LCR, CEA-LETI, CNET, GEC-Hirst Research Centre, and Cambridge University for providing the samples and collaboration which enabled this work to be performed.

REFERENCES

[1] M. Montier Esprit '86' Results and Achievements, pp 197-205.

[2] S. L. Partridge, IEDM '86, pp 428-430.

[3] e.g. papers in Proc. Symp. on Laser and Electron Beam Processing of Electronic Materials, eds. C. L. Anderson, G. K. Celler, G. A. Rozgonyi, vol. 80-1, Electrochem. Soc.

[4] J. R. Davis, R. A. McMahon, H. Ahmad J. Electrochem Soc. 132, pp 1919-1985.

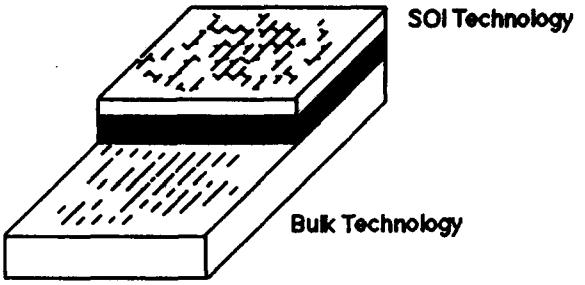
[5] A. J. Auberton-Herve, J. P. Joly et al ESSDERC 1984.

[6] M. Haond, D. Dutratre, D. Bensahel MRS Europe 1985.

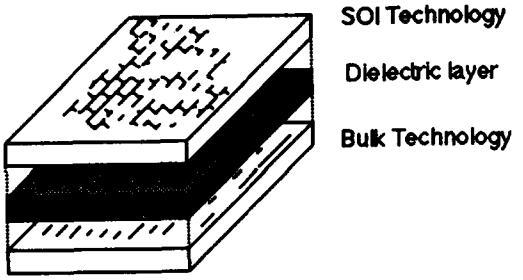
[7] W. A. Lane et al IEE colloquim on Intelligent Power Devices, March 1987.

[8] L. Hobbs, 1987, MEngSc. Thesis, University College Cork (not published).

[9] G. F. Hopper et al, Electronics Letters, Vol. 20, No. 12, pp 500, 1984.



(a)



(b)

FIGURE 1

(a) The mezzanine structure and (b) The stacked structure.

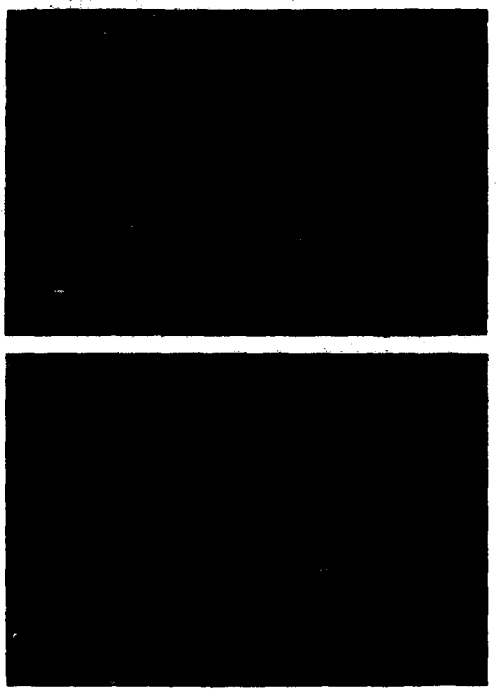


FIGURE 2

Characteristics of 3um channel length NMOS and PMOS transistors fabricated in e-beam recrystallised polysilicon.

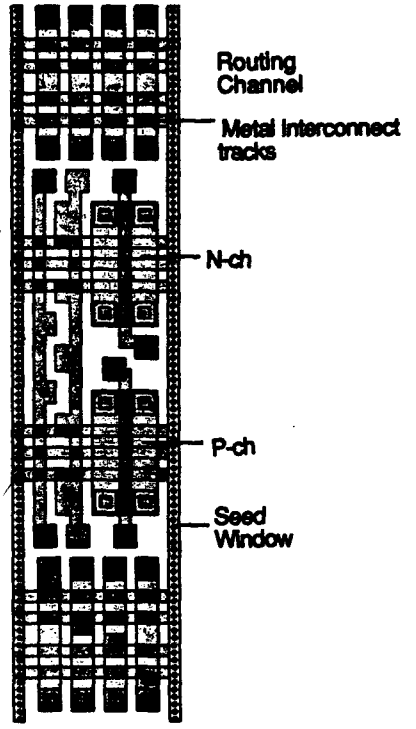


FIGURE 3

SOI CMOS Gate Array layout.



FIGURE 4

Polysilicon track showing mass transport effects.

PARAMETER	BEFORE ZMR	TYPICAL VALUES AFTER ZMR
VT [V]	[1.14 ± 0.26]	1.25
BVDS [V]	> 50	> 50
RON 5V @ VGS	33 ±13 Ω	38 Ω
15V	22 ±12 Ω	28 Ω

TABLE 1

DMOS characteristics before and after ZMR step.

ANALYSIS OF NONUNIFORMLY DOPED SOI MOSFET's

P. Paelinck*, O. Vancauwenberghe and F. Van de Wiele

Laboratoire de Microélectronique, Université Catholique de Louvain
Place du Levant 3
1348 Louvain-la-Neuve, Belgium

A one-dimensional device simulator is developed for nonuniformly doped SOI MOSFET's which allows to calculate accurately and reliably their electrical characteristics in the linear region. NMOS and PMOS transistors have been successfully optimized in relation to the process implantation parameters. Comparison with experimental results is presented.

1. INTRODUCTION

There is some difficulty in optimizing SOI transistors performances with classical simulation tools. This stems from the fact that conventional numerical device simulators are not adapted to the SOI-related characteristics of such devices : technological (nonuniform doping level, fixed oxide charges and interface states densities) and geometrical (film finite thickness, buried oxide thickness) parameters, presence of a fourth terminal (back gate contact).

We developed a simulator, coupled to process simulator SUPREM [1], which takes all these data into account. It is based on the one-dimensional solution of Poisson equation in the linear region. Further, the problem of equilibrium (flatband voltages) has been carefully treated in order to determine the built-in voltage caused by a varying doping level [2] that we adapt to a SOI structure. The computer code allows the current to be reliably calculated from the weak inversion criterium. Furthermore, its simplicity makes it a fast SOI-oriented tool for the development of successful device technology.

2. DISCUSSION

2.1. n-channel transistor

In SOI technologies, the backside SiO_2/Si interface is generally of rather poor quality, owing to the high fixed oxide charges and interface states densities [3]. In NMOS transistors, this results in the persistence of a significant leakage current [4] in the OFF state. This difficulty can be eliminated by applying a negative bias at the back gate terminal. Therefore, the back channel inversion is always impeded, provided that the back gate bias is appropriate. However, this has the drawback of requiring an additional external power supply and the SOI MOSFET becomes a four terminals device. An alternative solution consists in performing a deep boron implantation [5] near the backside interface in addition to the superficial one (Fig.1).

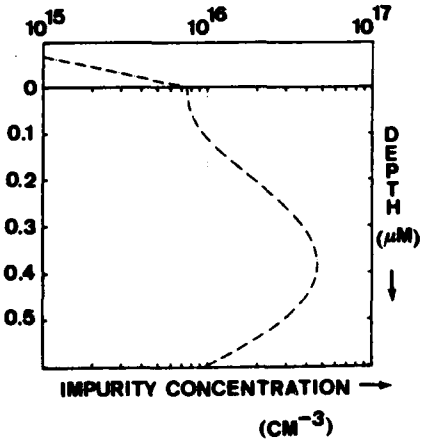


FIGURE 1

Net impurities concentration versus depth
E_{front}=40 keV, $\Phi_{front}=10^{11}$ cm^{-2}
E_{back}=160 keV, $\Phi_{back}=1.5\ 10^{11}$ cm^{-2}

* P. Paelinck is financially supported by I.R.S.I.A.

In this case, the SOI transistor comprises three terminals like a conventional MOS transistor, since the back gate terminal is left floating. The parameters of both implantations have to be carefully selected. Indeed, it is well known that the front gate threshold voltage V_{Tf} sharply varies when the film is fully depleted, but that it remains constant when the backside interface is accumulated [6]. That's why we determined the implantations parameters in order to accumulate the backside interface. Our simulator enables us to investigate the back gate threshold voltage V_{Tb} and current sensitivity to the deep boron implantation dose, while still achieving the correct V_{Tf}. Figure 2 depicts the variation of V_{Tb} with boron dose.

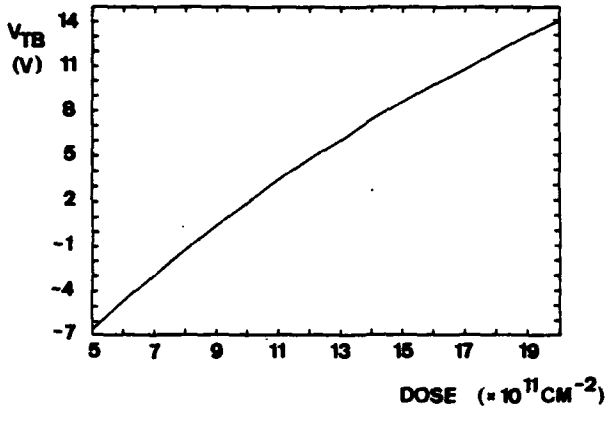


FIGURE 2

Back gate threshold voltage versus deep boron implantation dose (V_{Gb}=0 V)
Same technological parameters as in Fig. 1

As it may be expected, V_{Tb} increases with boron dose and becomes positive for doses over 9 10^{11} cm^{-2}. However, optimization can not be only based on threshold voltage control. Although V_{Tb} is positive, it is impossible to turn off the back channel leakage current for doses ranging from 9 to 13 10^{11} cm^{-2}. This is represented in Fig. 3 where the leakage current is plotted versus boron dose.

Further, the energy and dose of the superficial implantation have been chosen in such a way that the doping level near the front interface should not degrade the channel carriers mobility. We have also checked up that a possible lack of precision on the film thickness was of little consequence on the device electrical characteristics (Fig. 4).

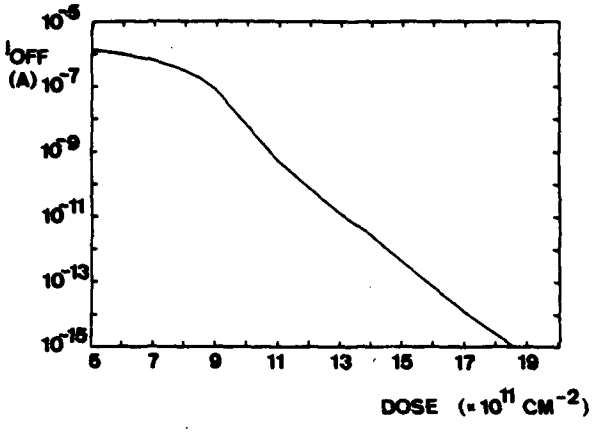


FIGURE 3

Back channel leakage current versus deep boron implantation dose (V_{Gb}=0 V).
Same technological parameters as in Fig. 1

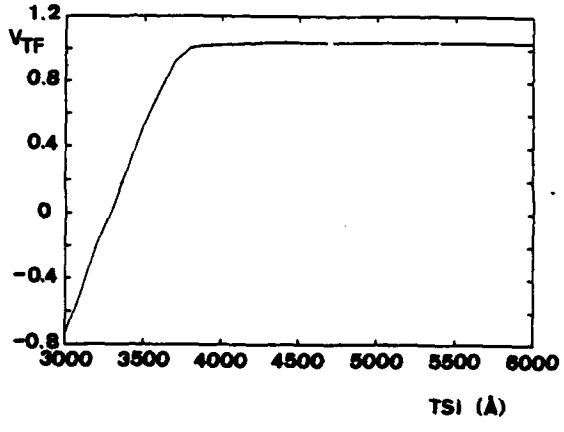


FIGURE 4

Front gate threshold voltage versus film thickness (V_{Gb}=0 V). Same technological parameters as in Fig.1

A comparison between simulated and experimental curves is presented in Fig. 5. Film thickness is 4800 Å and front and back oxide charges are respectively 10^{11} and 5 10^{11} cm^{-2}. It is

worth to note that the subthreshold current is correctly predicted by simulations.

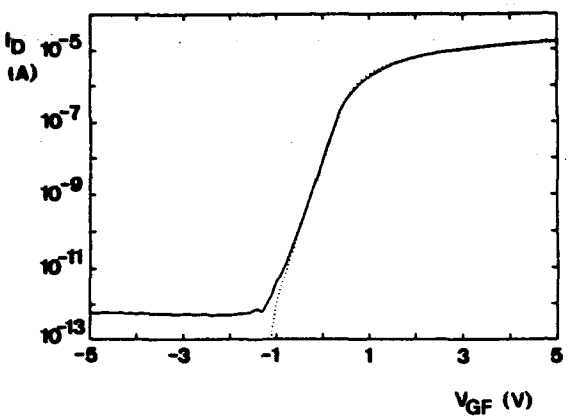


FIGURE 5

Experimental (___) and simulated (.....) (I_d, V_{Gf}) curves for NMOS transistor (V_{Gb}=0 V) W/L=20/8 microns

2.2. p-channel transistor

The optimization of PMOS transistors is quite different. The natural threshold voltage of the undoped active layer is too negative and the subsequent phosphorous implantation for substrate doping and threshold voltage adjustment reduces it still further. This problem can be circumvented in two ways : either implant only boron (fully depleted PMOS transistor) or counterdope the n-type substrate with a superficial boron implantation. Firstly we present the results we obtained for the latter.

Both phosphorous and boron implantations give rise to a shallow junction. Moreover, SUPREM simulations show that a second deeper junction may be induced above the backside interface; this can be explained by the fact that the film is initially intrinsic. We kept constant the phosphorous implantation parameters and studied the influence of boron implantation. According to the quantity of implanted boron ions, the film can be considered as a n,p-n,p-n-p or p-type substrate. The simulations allow to determine energies and doses in order to minimize junction leakage current and to adjust the buried-channel threshold voltage. Nevertheless, threshold voltage strongly depends on the implanted boron dose and Fig. 6 indicates a nearly linear dependence of V_T with the dose.

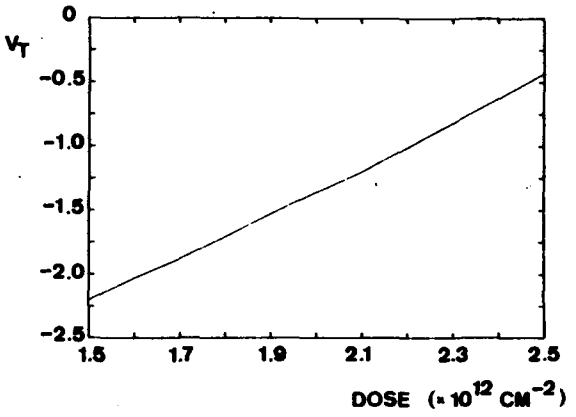


FIGURE 6

Threshold voltage versus boron dose for buried-channel PMOS transistor

In the second approach, the single boron implantation leads to a fully depleted PMOS transistor. The basic idea is that electrons accumulate at the backside interface when the back gate threshold voltage is negative, and the depletion zone extends to the front interface. It is thus important to avoid any hole conduction in volume. This can be achieved providing that V_{Tb} is sufficiently negative. Therefore, the key parameters are the impurities concentration near the backside Si/SiO_2 interface, the buried oxide thickness and the fixed buried oxide charge density. The influence of the latter appears in Fig. 7 It shows that the lesser quality of the back interface tends to stabilize the front channel threshold voltage. This is due to the fact that the buried oxide is thick, so that V_{Tb} is negative and allows electrons accumulation. Moreover, the depletion zone depth is limited by the film finite thickness. Good agreement has been found between simulation and experience (Fig. 8).

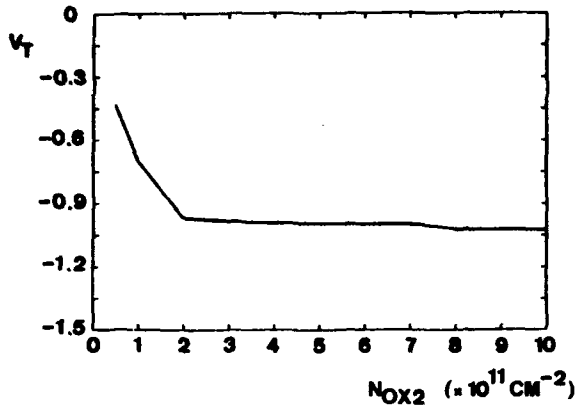


FIGURE 7

Threshold voltage versus fixed back oxide charge density (V_{Gb}=0 V) for fully depleted PMOS transistor

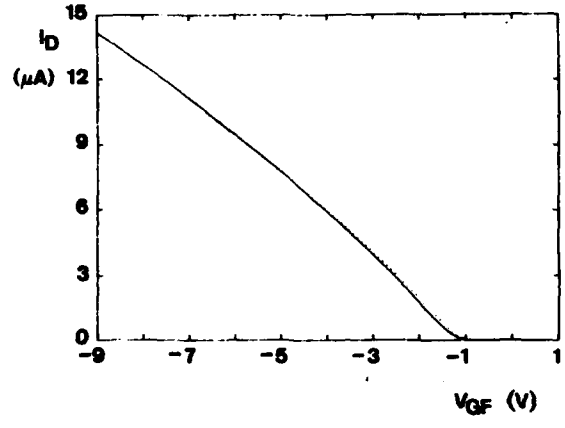


FIGURE 8

Experimental (___) and simulated (.....) (I_d,V_{Gf}) curves for PMOS transistor (V_{Gb}=0 V) W/L=20/6 microns

3. CONCLUSIONS

In conclusion, we have developed a one-dimensional numerical device simulator for nonuniformly doped SOI MOSFET's. It also includes geometrical (film and buried oxide thicknesses) and technological (fixed oxide and interface states densities) device parameters. Therefrom, the influence of both superficial and deep boron implantations parameters on the electrical characteristics of NMOS transistors has been clarified in order to get the correct threshold voltage and suppress the back-gate-related parasitic leakage current. Two types of PMOS transistors (buried-channel and fully depleted) have been investigated and reliably optimized. Their sensitivity to implantation parameters has also been studied.

REFERENCES

[1] Antoniadis, D.A., Hansen, S.E. and Dutton, R.W., SUPREM II - A program for IC process modeling and simulation (SEL-78-020 Stanford Electronics Labs, Stanford University , 1978)

[2] Van de Wiele, F. Solid-State Electronics (1984) 824

[3] Le, H.P. and Lam, H.W., IEEE EDL (1982) 161

[4] Lam, H.W., Sobczak, Z.P., Pinizzotto, R.F. and Tasch Jr, A.F., IEEE TED (1982) 389

[5] Mc Greivy, D.J., IEEE TED (1977) 730

[6] Lim, H.K. and Fossum, J.G., IEEE TED (1983) 1244

Session C2.1

GaAs Device Modelling

Chairman: E. Munoz Merino

Tuesday, September 15, 1987

MODELLING OF THE DRAIN LAG EFFECT IN GaAs MESFET's AND ITS IMPACT ON DIGITAL IC's

Thierry DUCOURANT, Marc ROCCHI

LEP : Laboratoires d'Electronique et de Physique Appliquée
Membre de l'Organisation de Recherche Internationale de Philips
3, avenue Descartes, 94451 LIMEIL-BREVANNES CEDEX, France

1. INTRODUCTION

N channel GaAs MESFET's directly implanted into semi-insulating substrates exhibit many low frequency current drifts which degrade the performances of GaAs digital IC's and result in comparator hysteresis, memory cell unstabilities, or abnormally low output buffer levels.

They all can be attributed to the same "drain lag" effet (1, 2). Though it is now well establisned that this effect is mainly caused by the low frequency response of the interface between the active channel and the semi-insulating substrate, its exact mechanism has not been completely clarified yet. Based on extensive measurements either in the time or the frequency domain, we propose a comprehensive model of the drain lag effect, to be included in CAD models for GaAs MESFET's.

2. BASICS OF DRAIN LAG EFFECT

The "drain lag" effect is related to the excess drain current of a MESFET resulting from a fast drain-to-source voltage transient. It also corresponds to the variation versus frequency of the output conductance (g_d). As a matter of fact, this effect results from the combination of two related phenomena : a frequency dependent current injection into the substrate (3) and a related back modulation of the channel (4). The injection results from the distortion of the electric field lines in the channel due to the formation of a high field domain at the drain side and from the lowered substrate resistivity under these high field conditions (figure 1). As electrons are injected, some of them get trapped by the deep centers located in the substrate near the interface which modifies the space-charge thickness in the substrate and consequently in

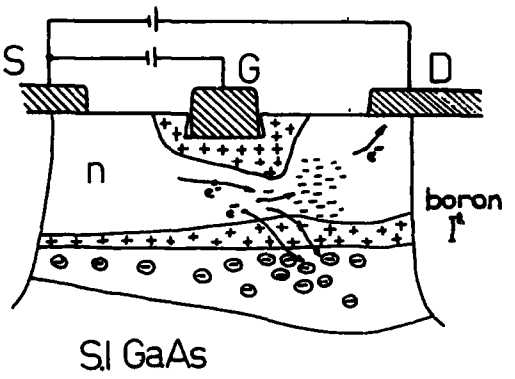


Figure 1

the channel. This results in frequency dependent electron screening, which limits the injection and leads to a back modulation of the channel.

As a conclusion, our physical analysis shows that in addition to being frequency dependent, the drain lag effect also strongly depends on V_{GS} and V_{DS} which fix the electric field conditions at the interface.

3. CAD MODEL OF DRAIN LAG EFFET

Though the frequency dependence of the drain to source impedance is complex, it can be simply fitted with one single pole at 100 Hz (measured data). For most applications only DC

(bias points) and high frequency operation have effectively to be considered, and the single pole approximation can then be used.

We propose the equivalent FET model of figure 2, associated with the following equations for the injection current I_{inj} and the channel current I_{ch} :

$$I_{inj} = \left(\frac{(V_{dsi}+(\alpha - 1)\ V_r)}{R\,[V_{gsi} - V_t]}\right).\left(\frac{V_{dsi}-(V_{dsi})_{sat}}{V_{dsi}+(V_{dsi})_{sat}}\right)$$

for $V_{dsi} \geq (V_{dsi})_{sat}$

$I_{inj} = 0$ for $V_{dsi} < (V_{dsi})_{sat}$

$I_{ch} = G_o\ f\ (V_{gs} - \beta\ V_c)\ g\ (V_{ds})_{sat}$

$R\ [V_{gsi} - V_t]$ characterises the injection resistance which depends on the intrinsic gate to source voltage V_{gsi}; $(V_{dsi})_{sat}$ is the intrinsic drain to source saturation voltage, V_t the threshold voltage.

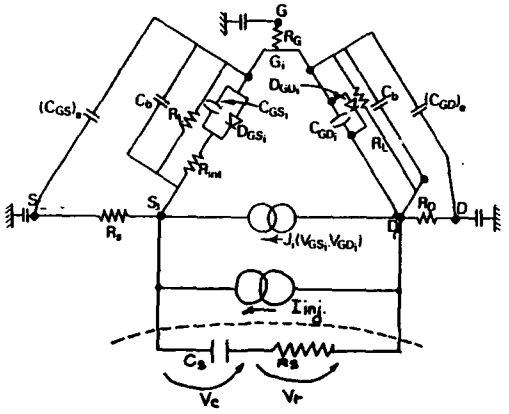


Figure 2

Equivalent CAD model of the MESFET

The coefficient α is a first order approximation of the ratio (α') of the high frequency g_d ($\geq$ 10 MHz) to the low frequency g_0 ($\leq$ 0.1 Hz). The small difference between α and α' is due both to the feedback effect of the injected current through the access resistances and to the back modulation factor β.

Ideally, α' should be equal to 1. It has been measured at V_{ds} = 1.5 V for various substrates between 0.1 Hz and 10 MHz. As expected, the module of G_D increases with frequency while its phase reaches a maximum of approximatively 10° around 100 Hz. Figure 3 is a particular example showing that the measured ratio α' is around 3 for V_{gs} bias ranging from 0 V up to .7 V. This agrees well with the computed data for α = 3 and β = 0.01.

Figure 4 shows the computed I-V curves at 0.1 Hz (DC), 100 Hz (curve tracer). The hysteresis loop clearly appears, as well as the frequency dependent available I_{DS} current.

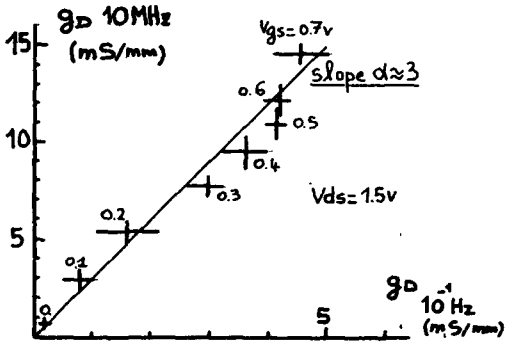


Figure 3

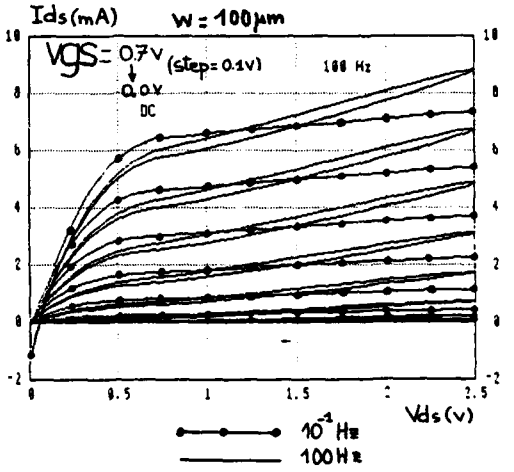


Figure 4

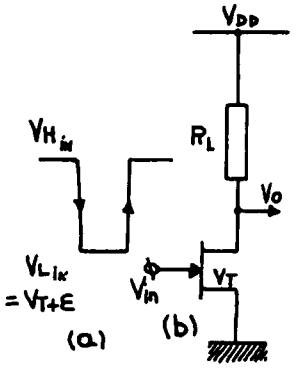
I-V curves

4. SIMULATION OF DIGITAL CIRCUITS

The effect of the drain lag can be investigated for digital circuit design.

The main conclusions are :

- the impact remains low for a basic DCFL inverter (figure 5) as long as the output voltage is clamped to .7 V by the next inverter (low field conditions).
- when the full V_{ds} swing is needed (output buffers, CML logic), the high output logic level first switches, then lags to its final value (figure 5).



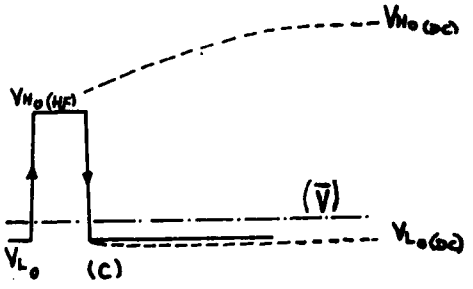


Figure 5 (a),(b),(c)

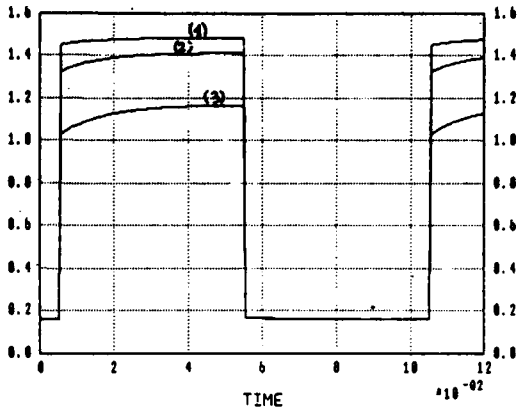


Figure 5 (d)

1. $(V_L)_{in} = V_T$
2. $(V_L)_{in} = V_T + 100$ mV
3. $(V_L)_{in} = V_T + 200$ mV

The transient high level (VM/hf) can be estimated analytically with a simple RC model for the drain lag effect (figure 6). Under these conditions, $\overline{V}$ is the mean voltage stored on the parasitic capacitor C_f. Moreover :

$$R_f\left[\overline{V}gs-Vt\right] = \frac{1}{gd_{dc}\left[\overline{V}gs-Vt\right]} = \frac{\alpha}{gd_{nf}\left[\overline{V}gs-Vt\right]}$$

This leads to :

$$V_h(Hf) = \frac{V_{dd}}{1+\alpha R_l/R_f} + \frac{\overline{V}(\alpha-1)}{\alpha + R_f/R_l}$$

$\overline{V}$ is directly related to the duty cycle. A worst case for a low to high transition is $\overline{V}$ = 0 (duty cycle $\ll$).

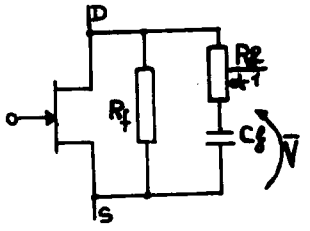


Figure 6
Simplified R-C model

A lagging percentage η_1 can be expressed as :

$$\eta_1 = 1 - \frac{1 + R1/Rf\ [Vgs_Vt]}{1 + \alpha(R1/Rf\ [Vgs-Vt]}$$

Rf [Vgs] being derived from low frequency measurements, the following data can be computed :

$\mathcal{E} = 0 \rightarrow \eta_1 = 1\ \%$, $\mathcal{E} = 0.2\ V \rightarrow \eta_1 > 10\ \%$.

The coefficient η_1 has been measured and computed with the complete CAD model. The clock period was 1 s for a duty cycle of 10^{-6}. Figure 7 demonstrates a reasonably good agreement between estimated, measured and computed data. It is clear that a buffer follower stage has to be added to the gate to improve the low logic level (buffered DCFL) ; moreover high value V_t should be used ($V_t \geq 150$ mV).

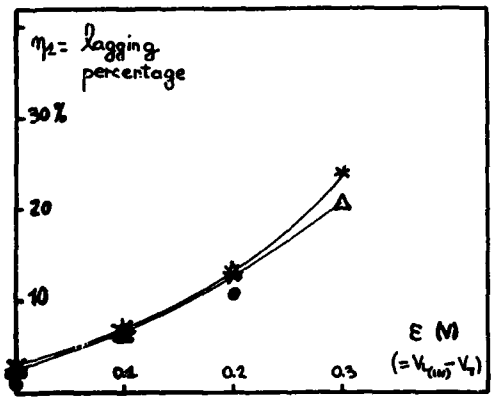


Figure 7

* measured

Δ computed (complete CAD model)

● computed (simple R-C model)

REFERENCES

(1) M. Rocchi, "Status of the surface and bulk parasitic effects limiting the performances of GaAs IC's, ESSDERC 84'

(2) T. Ducourant, "3 GHz, 150 mW 4 bit GaAs ADC, GaAs Symposium 86"

(3) Eastman & Shur, "Substrate current in GaAs MESFET's, IEEE Trans. on ED, september 79

(4) S. Makram-Ebeid, "The roles of the surface and bulk of the SI substrate in LF anomalies of GaAs IC's, IEEE Trans. ED, March 85.

MODELLING AND SIMULATION OF WAVE PROPAGATION EFFECTS IN MESFET DEVICES BASED ON PHYSICAL MODELS

Giovanni GHIONE, Carlo U. NALDI,

Dipartimento di Elettronica, Politecnico di Torino,
Corso Duca degli Abruzzi 24, 10129 TORINO, Italy

Propagation effects along the electrodes of single- and multi-gate MESFETs are analyzed by means of new model which directly exploits two-dimensional small-signal simulation. The electromagnetic behaviour is characterized through a quasi-TEM multiconductor line model, and preliminary results are presented concerning the influence of gate width on the overall device performances.

1 INTRODUCTION

As well known, propagation along the electrodes affects the small-signal performances of MESFET devices at microwave frequencies, and plays an important role in further deteriorating their gain. An accurate analysis of such effects is therefore essential both to achieve better small-signal models for the device and to gain further insight into propagation phenomena so as to reduce their influence by means of proper design rules. Finally, the analysis of propagation is relevant to understanding the operation of travelling-wave MESFET amplifiers.

The simulation of propagation effects in MESFETs was addressed during past years through a variety of approximate techniques, based on modelling the device as a transmission line directed along the gate electrodes, whose parameters are derived partly from electromagnetic analysis and partly from small-signal lumped MESFET equivalent circuits. In early works [2] single-gate MESFET were considered and propagation along drain and source fingers was neglected; moreover, ohmic losses in the latter were not accounted for. These simplifying assumptions have been shown [6] to lead to substantial errors in many practical cases. Later, single-gate MESFET were modelled as a three-conductor transmission line [3], and finally, attempts were made to model multi-gate MESFETs [4]. More recently, this simple quasi-TEM transmission line approach to the problem was questioned and a full-wave analysis by means of mode matching techniques was proposed which, owing to its complexity and CPU intensity, is confined to considering single gate MESFETs and uses a rather crude model for the active region [5,6]. Nevertheless, this analysis clearly suggests that the propagation modes supported by MESFET structures are quasi-TEM, and that ohmic losses both within the active region and on the metallizations cannot be neglected. Hence, a lossy transmission line equivalent circuit should yield a good electromagnetic model. This conclusion is confirmed in [7], where it is shown how a simple, computationally unexpensive, quasi-TEM model is in excellent agreement with full-wave analysis.

In spite of the progress achieved in understanding the electromagnetic behaviour of microwave MESFETs, the small-signal model of the active region employed in distributed device simulation was often approximate. The aim of the present paper is to achieve an accurate and flexible characterization of distributed effects in MESFETs by coupling a model of the active region directly derived from small-signal device simulation based on physical models [1] to a multiconductor transmission-line model including all relevant electromagnetic phenomena (electrode losses, effect of both external and internal capacitive and inductive coupling between electrodes, effect of substrate).

2 THE DISTRIBUTED MESFET MODEL

In order to define the problem, let us consider a two-gate power MESFET with source air-bridge whose cross-section is shown in Fig.1. Although all active phenomena are confined to the dashed (active) regions, the (passive) rest of the device cannot be neglected in modelling microwave small-signal operation. In fact, source, drain and gates are both capacitively and inductively coupled both through the substrate and the air; moreover, ohmic losses are present in all metallizations, and particularly in gate fingers, and an internal distributed inductance (though often negligible) is associated to each electrode. In order to define the elementary cell of the four-conductor transmission line equivalent to a MESFET section, we must introduce the per-unit-lenght (p.u.l.) admittance matrix and impedance matrix of the line. In the case at hand, such matrices read:

$$(1a)\quad Y_{ij} = Y_{a1ij} + Y_{a2ij} + j\omega(C_{1ij}+C_{2ij})$$
$$(1b)\quad Z_{ij} = j\omega L_{ij} + R_{ij}(\omega)$$

where Y_{akij} is the p.u.l. admittance matrix of the k-th active region (here k=1,2), C_{1ij} is the internal capacitance matrix accounting for coupling within the substrate outside the active regions, while C_{2ij} is the external (air) capacitance matrix. The inductance matrix L_{ij} can be approximately evaluated in terms of the inverse of the capacitance matrix in vacuo; the internal inductance of wires can oe included, if significant. Finally, R_{ij} is the resistance matrix of the electrodes. The circuit interpretation of (1a) and (1b) is shown in Fig.2 . In regard to the evaluation of the passive contributions to Y_{ij}, well known techniques are employed to compute L_{ij} and C_{ij}, such as spectral-domain Green's function techniques [9], which are well suited when the FET is planar. Complex geometries with non-planar dielectric layers require the introduction of more flexible finite-element (FEM) techniques [10]. The frequency dependent resistance matrix is evaluated as in [11] accounting for non-uniform current density distribution within the conductor cross section.

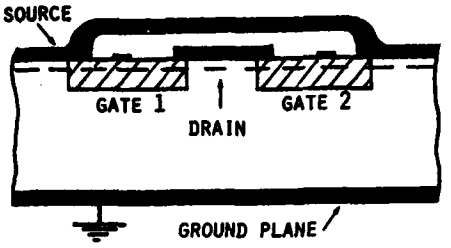


Fig. 1

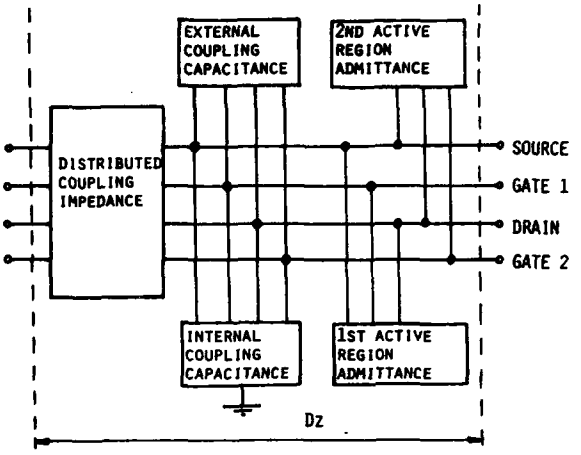


Fig. 2

Concerning the active part of the device, the relevant admittance matrices are directly obtained from a two-dimensional MESFET simulator (MESS, [8]). The steady-state analysis is performed by means of full Newton techniques, while small-signal characterization is obtained by Fourier transforming the time-domain response of the linearized Poisson and continuity equation. The Scharfetter-Gummel scheme on a triangular grid is used for discretizing the continuity equation, while Poisson's equation is treated by FEM with charge lumping. Both planar and recessed-gate devices, with arbitrary doping profile, are simulated.

From the knowledge of the p.u.l. admittance and impedance matrix, the generalized Kirchhoff

equations for the line voltage $\underline{v}(z)$ and current vectors $\underline{i}(z)$ read:

$$(3a) \qquad \partial \underline{v}(z)/\partial z = -Z\underline{i}(z)$$

$$(3b) \qquad \partial \underline{i}(z)/\partial z = -Y\underline{v}(z)$$

and the propagation constants of the lines $k_i = \beta_i - j\alpha_i$ are solution to the eigenvalue problem:

$$(4) \qquad [\, k^2 I - ZY \,]\, M_v = 0$$

where I is the identity matrix, M_v the voltage eigenvector matrix. From M_v and the corresponding current eigenvector matrix M_i the characteristic impedance matrix of the line can be computed and the scattering matrix of a multiconductor line of finite length is derived by means of standard techniques (*cfr.* [12]).

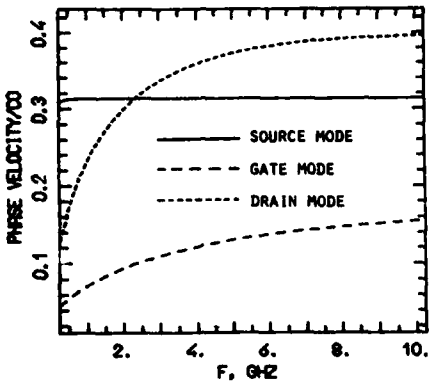


Fig.3a - Phase velocity

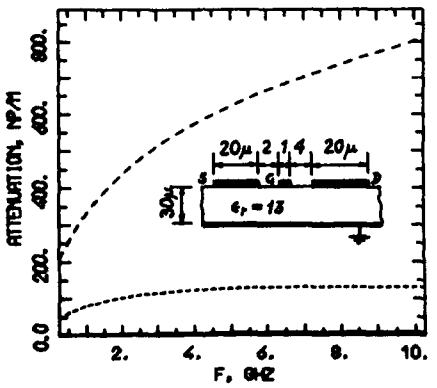


Fig.3b - Attenuation

3 RESULTS

Propagation modes in multiconductor MESFET structures can be roughly divided into three classes: gate, drain and bulk modes [6]. Bulk modes are microstrip-like modes with low losses and almost constant phase velocity; the all lines are driven at the same potential and there is no gain mechanism. Drain modes are basically high-loss modes whose voltage distribution is odd with respect to gate and source lines. Finally, gate modes show very high losses with unbalanced voltage distribution and built-in gain mechanism. The analysis confirms the remarks made in [6] according to which MESFET fingers do not actually support growing waves as in devices with bulk gain. Actually, even in the limiting case wherein gate and drain fingers are lossless and coupled only by the transconductance, the current profile on the drain electrode would not be exponential, but only linear. This is basically due to the fact that the gain mechanism of MESFETs acts by

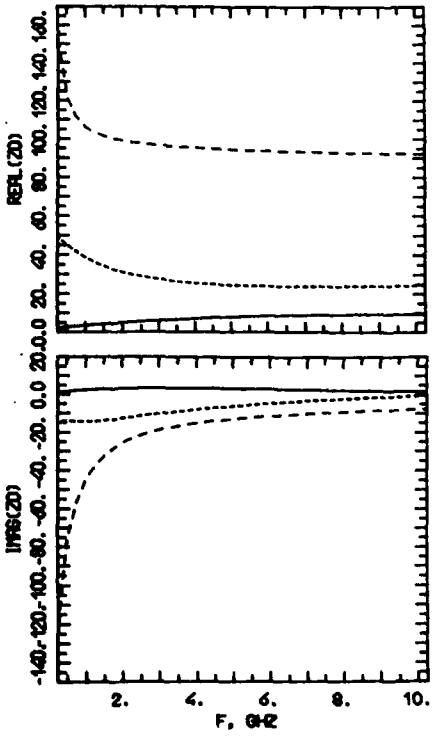


Fig.3c,d - Modal impedances

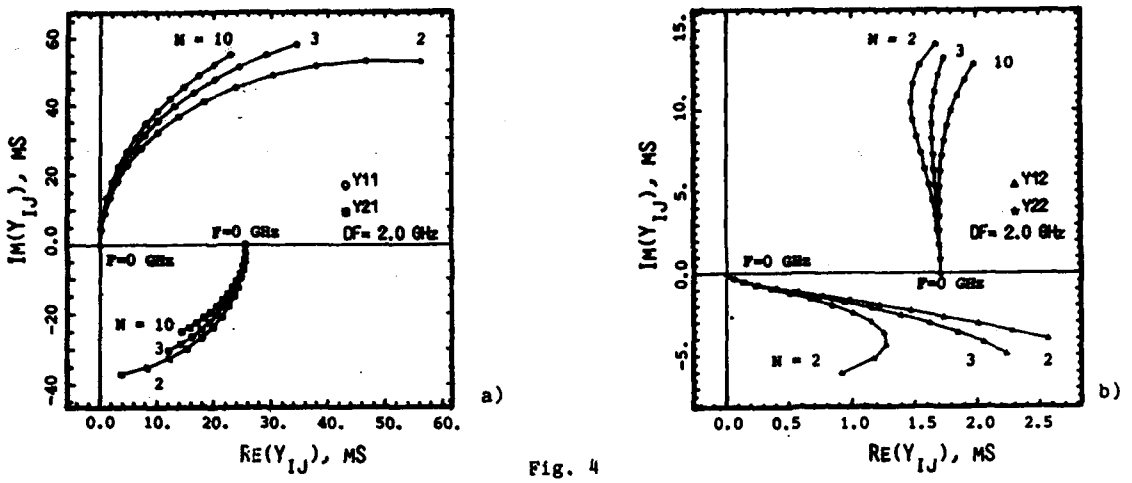


Fig. 4

shunt injection of current, i.e. is perpendicular to the propagation direction. Another point which deserves attention is the influence of ohmic losses on gate modes. Owing to the reduced cross-section of gate, this electrode has very high p.u.l. resistance (up to several KΩ/m), unless special structures are exploited (e.g. gate air-bridge FETs). As an example, the phase velocity, attenuation and complex characteristic impedance are shown in Figs. 3a-d for a 1 μm single-gate MESFET; the other geometrical parameters are shown in Fig.3b.

In order to highlight the impact of distributed effects on small-signal parameter evaluation, let us consider the frequency behaviour of the admittance matrix of the same device of Fig.3b when the overall gate periphery W is kept constant and equal to 300 μm, while the number of gate fingers N, and therefore the length of each finger w, is changed. In Figg.s 4a, 4b the Y parameters for frequencies ranging from 0 to 28 GHz and for several values of N. Parasitic effects due to propagation are more dramatic beyond 10 GHz, but their magnitude is not negligible even at lower frequencies.

4 CONCLUSIONS

A quasi-TEM analysis of propagation effects in MESFET devices has been presented, which directly exploits two-dimensional small-signal analysis of the active region. Preliminary results on propagation characteristics are presented, and an example of how propagation effects can affect the small-signal device parameters is given.

ACKNOWLEDGEMENTS

This work has been partly supported by TELETTRA

REFERENCES

[1] Laux, S.E., IEEE Trans. ED, 1985(10) 2028

[2] Kuvas, R.L., IEEE Trans. ED, 1980(6) 1193

[3] Ren,Y.A., Hartnagel,H.L., Int. J. Electronics, 1981(5) 663

[4] Oxley,C.H., Holden,A.J., IEE Proc. Pt.H, 1986 (5) 335

[5] Heinrich,W., Hartnagel,H.L., Int. J. Electronics, 1985(4) 613

[6] Heinrich,W., Hartnagel,H.L., IEEE Trans MTT, 1987(1) 1

[7] Heinrich,W., IEEE Trans MTT, 1987(5) 487

[8] Ghione,G., Naldi,C., 3rd Meet. of GaAs Simulation Group, Duisburg, Oct. 1986.

[9] Ghione,G., Naldi,C., VI Riun. Naz. Elettromagn. Appl., Trieste, Oct. 1986, 75

[10] Cottrel,P.E., Buturla,E.M., IBM J. Res. Develop., 1985(3) 277

[11] Waldow,P., Wolff,I., IEEE Trans MTT, 1985(10) 1076

[12] Siegl,J., Tulaja,R., Hoffmann,R., Siemens Forsch.-u. Entwickl.-Ber., 1981(10) 228

ACCURATE NONLINEAR CHARACTERIZATION AND MODELING OF THE GaAs FET

Y. BONNAIRE and E. ALLAMANDO

Centre Hyperfréquences et Semiconducteurs - U.A. 287 C.N.R.S.
Université des Sciences et Techniques de Lille Flandres Artois
59655 VILLENEUVE D'ASCQ CEDEX - FRANCE

Contrary to previous conventional characterizations and modelisations, which employ the dc measurements of the drain current Ids(Vgs, Vds), we propose a new mode of nonlinear characterization of the GaAs FET. This is based on the knowledge of the voltage dependence of the transconductance gm and the output conductance gd values measured in the microwave frequency range. This new mode of characterization is thus more accurate and precise than conventional ones, because the latter are tainted with errors due to thermal and trap effects, which are parasitic with respect to the microwave behaviour of the FET. Thus, a numerical integration of the measured microwave parameters gives us the accurate quasi-static characteristics. On the other hand, we present an accurate model which takes into account the device nonlinearities.
Finally our method has been used to model a medium power amplifier, in X-band operating.

I. INTRODUCTION

Conception and optimization of microwave circuits, using the GaAs FET, require an accurate characterization of the device, especially in large signal and microwave operating. Moreover, it would be very interesting to be able to modelize it in the simplest way in order to reduce the computational time for use with conception and optimization of both devices and microwave circuits (CAO).

Many models have already been suggested :

a) Physical models using the basic equations of the FET are certainly the most accurate and have been very useful to understand the device operation. But the point is they are not easily suitable for CAO applications.

b) Phenomenological models consisting in describing as perfectly as possible the dc measured evolution of the drain current Ids as an analytical function of the gate-source Vgs and drain-source Vds bias voltages [1].

According to us, the latter mode of characterization is however not satisfactory enough and this for several main reasons :

- dc measurements are often imprecise ;
- Thermal and trap effects keep us from obtaining the microwave characteristics ;
- The knowledge of the analytical function Ids (Vgs, Vds) does not necessarily imply the validity of the first partial derivates of the drain current which are the dynamic parameters gm and gd ;
- These are indeed the fundamental parameters we have to measure accurately (specially gm), because the microwave behaviour of the FET is strongly dependent on their microwave values, and on their voltage evolutions, characteristic of the intrinsic nonlinearities.

We propose to show that it is possible to obtain the quasi-static characteristics, available in microwave, from the RF measurements of the transconductance gm over a large range of bias voltages.

On the other hand, we suggest quite simple phenomenological expressions able to describe the evolution of gm, and then of the drain current as functions of the bias voltages Vgs and Vds.

Finally, we present a theoretical modeling using those expressions, that we will valid by means of the study of a medium power amplifier, in X-band operating.

II. ACCURATE NONLINEAR CHARACTERIZATION

A. Principle

By using the microwave measurements of the first partial derivate of the drain current with respect to the gate-source voltage, we mean gm, we can deduce the quasi-static characteristics by a numerical integration :

$$Ids(Vgs_i)|_{Vds} = \sum_{k=1}^{i} gm(Vgs_k)\ \Delta Vgs_k$$

With Vgs_1 = Pinch-off voltage.

Remark that we decided to use the gm measurements rather than gd ones because we consider that gm is the most important parameter, with regard to microwave behaviour and nonlinear properties of the device.

Furthermore, by a numerical derivation, we may infer from those characteristics the output conductance

gd(Vgs, Vds) = ∂Ids/∂Vds

and compare it with the experimental results in order to valid the method self-consistence.

B. Experimental set-up

In practice, the measurement of the intrinsic transconductance gm is accurately made at moderately high frequency range (2-4 GHz) by using a microwave network analyser (HP 8510). After substraction of access elements, determined from measurements at Vds = 0, we deduce the intrinsic elements of the FET equivalent circuit [2]. That method, which is very fast and doesn't need long computational time, is very suitable for the determination of the parameters gm and gd over a wide range of bias voltages, corresponding to the usual operating of the device. We so determined the parameter gm and gd values every other 0,1 v for Vgs and every other 0,2 v for Vds to be able to get an accurate integration.

C. Results

For example, we show Fig. 1, the quasi-static characteristics obtained from experimental evolutions of gm (Fig. 2) for a commercial submicrometer FET NE 673 (gate length is Lg ≈ 0,3 µm).

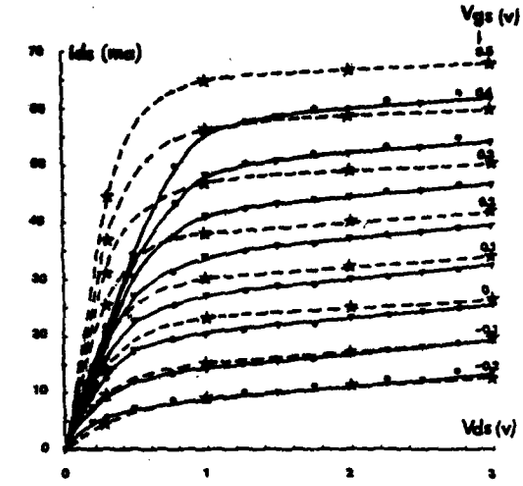


Fig. 1 : Output characteristics Ids(Vgs,Vds) for the NE 673.
★ — ★ — dc measured
▽ — ▽ — quasi-static deduced from experimental measurements.
● theoretical given by our model

The comparison of those characteristics with the dc measurement results (Fig. 1) exhibits important differences between the general shapes of the curves specially for high drain currents, all that tends to prove the non accuracy of dc measurements.

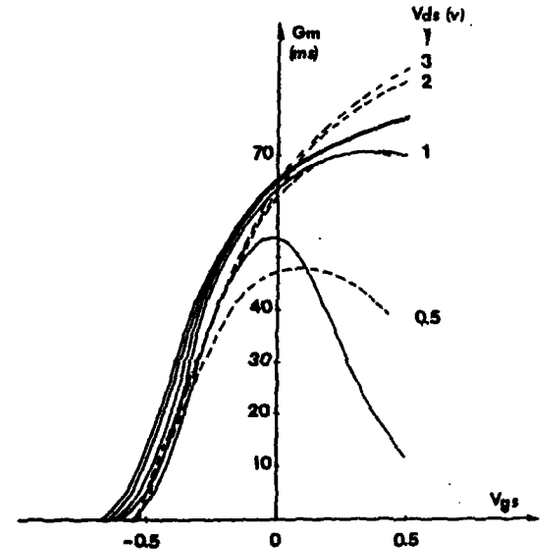


Fig. 2 : Bias dependence of gm.
——— measured
– – – – computed (expression (1) with α=1.18 ; Vo = 0.5 ; Vp = -0.65 ; a=0.46 b = 0.37.

III. MODELING OF THE FET

A. Equivalent circuit

According to the simplified large signal model shown Fig. 3, we consider that the intrinsic parameters, determined from small signal measurements become nonlinear if we consider their evolutions over a wide range of bias voltages [3].

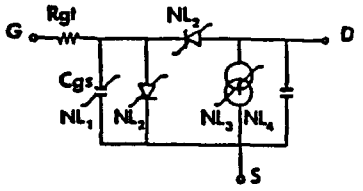


Fig. 3 : Model of the FET used in simulation.

B. Nonlinear analytical representation

We propose a new nonlinear approach based on an original analytical representation of the nonlinear voltage dependence of the parameter gm. Thus, the measured evolution of the transconductance gm versus Vgs and Vds given fig. 2, leads us to employ the phenomenological expression :

$$gm(Vgs,Vds)=gmo(Vds).th[\alpha(Vgs+Vo)/Vp].th(Vds/Vs) \quad (1)$$

By mathematical integration, we may obtain a new phenomenological expression of the quasi-static characteristics Ids(Vgs, Vds).
Indeed :

$$Ids(Vgs, Vds) = \int_{-Vp}^{Vgs} gm(Vg's, Vds)\, dVg's \quad (2)$$

$$= \{Iso.Ln[ch(\alpha \frac{Vgs+Vo}{Vp})]+gdo.Vds\}.th(\frac{Vds}{Vs}) \quad (3)$$

with $Iso = gmo(Vds).Vp/\alpha$

Where the drain-source saturation voltage Vs is described by a linear dependence with the bias voltage Vgs : $Vs = a.Vgs+b$.

Moreover, gdo corresponds to the FET output conductance in saturation mode. gdo is looked upon as independant from the bias voltage Vgs, for it exhibits very small variations as it can be seen from experimental measurements (Fig. 4).

This model also requires four parameters : α, gmo, a and b, that are optimized to provide the best average fit of the theoretical to the measured evolutions of gm.

Furthermore, we consider two other nonlinearities :

- NL1 : Voltage dependence of the gate-source capacitor classically described by :

$$Cgs = Co.(1 - Vgs/\Phi)^{-\frac{1}{2}}$$

- NL2 : Direct current of the schottky gate junction into source or drain circuit.

C. Validity

Experimental and theoretical voltage evolutions of gm for a commercial submicrometer FET are firstly compared Fig. 2. Secondly, we also compared Fig. 1 the corresponding evolutions of the drain current. In those two cases, we can notice a good agreement between experimental results and values given by the phenomenological expressions hereby proposed (1) (3), which can be considered as valid.

Finally, the consistence of our method is proved Fig. 4, by comparison between gd measured and computed evolutions.

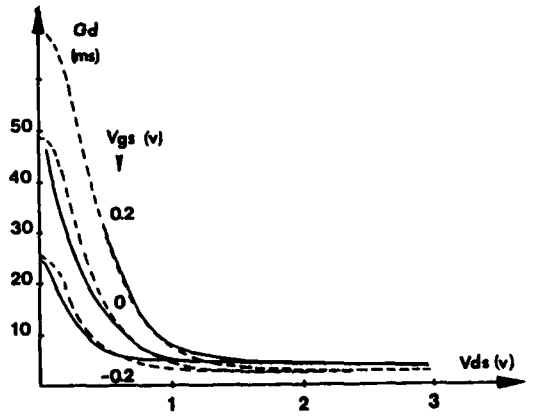


Fig. 4 : Bias dependence of the output conductance gd.
——— measured — — — calculated.

D. Calculation

The numerical treatment of this model consists [4], at each time, in describing the instantaneous value of the gate-source voltage. Further, by using the model proposed, with those nonlinear representations, we determine the instantaneous values of both the drain-source voltage and the drain current. For this, we take into account the drain circuit termination. Finally, a Fourier expansion gives the power values of the frequency components.

IV. VALIDATION

Our new phenomenological expressions proposed and our model have been validated in the case of the study of a GaAs FET used as a microwave medium power amplifier in X-band. Indeed, good agreement between theoretical and experimental results is obtained (Fig. 5) and small signal gain as well as saturation power are successfully simulated.

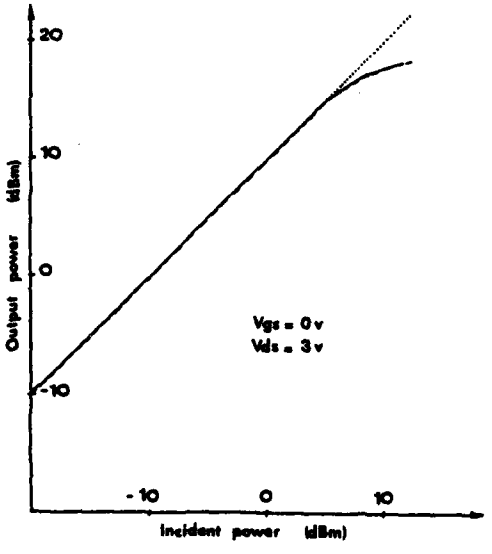


Fig. 5 : Output versus input power at 10 GHz. (transistor NE 673).
——— Measured
– – – – Computed

V. CONCLUSION

Contrary to conventional modelizations, which are based on dc measurements, we suggest an accurate characterization valuable for both nonlinear dependence and microwave behaviour of the GaAs FET. Our model, which will enable the description of the microwave behaviour of the FET used in great signal condition, has been confirmed by comparison between theoretical prediction and experimental results for a commercial submicrometer FET.

Our model efficiency results from the use of new and simple analytical description of the characteristics, directly issued from the precise RF measurements. The self consistence of our method, as well as the good agreement observed between measurement and prediction make us fell that it could profitably be used for CAO applications.

ACKNOWLEDGMENT

The authors wishe to thank Prof. Y. LEROY for his continuous encouragement and helpful discussions.

REFERENCES

[1] W.R. CURTICE, "A MESFET model for use in the design of GaAs integrated circuits", IEEE Trans. on Microwave Theory Tech., Vol. MTT-28, pp. 448-456, May 1980.

[2] A. CAPPY, "Propriétés physiques et performances potentielles des composants submicroniques à effet de champ". Doctorat d'Etat Univ. Lille I, Déc. 1986.

[3] R. MINASIAN, "Large signal GaAs mesfet model and distorsion analysis". Electron. Lett., 1978, 14, pp. 183-185.

[4] E. ALLAMANDO, N.E. RADHY, E. CONSTANT, "Broadband high-order frequency multipliers by using dual-gate Mesfet's". Proc. 15th European Microwave Conference, Paris, Sept. 1985.

A THREE-DIMENSIONAL MODEL WITH DISTRIBUTED ELEMENTS FOR GAAS MESFETS AND SIMILAR DEVICES

W.Wiesbeck, S. Haffa and H. P. Feldle

University of Karlsruhe
7500 Karlsruhe Germany

The technological advances in active microwave and mm-wave semiconductors like GaAs-MESFETs and HEMTs e.g. planar and recessed gate structures will result in the near future in components with cutoff frequencies of several hundred GHz. For modelling of these elements it is mandatory to take distributed elements and active and passive coupling into account. To realize this, a three-dimensional model has been developed, which involves distributed elements, coupling and wave propagation. The results show the typical effects of a slow and a fast wave propagation in lossy media with open-ended lines.

1. INTRODUCTION

Modelling of semiconductor devices has gained much interest in the last years, as it is shown by numerous publications. The modelling results are given in the form of S-parameters fairly accurately for CAD of microwave and mm-wave circuits. The standard models for GaAs MESFETs are based on a six element equivalent circuit. The inner transistor two-port is surrounded by parasitic elements of the chip, the case and so on, which in turn form two-ports.[1] Models based on this equivalent circuit, or similar circuits, are restricted to frequencies for which lumped element equivalents are valid, not more than 18 GHz to 20 GHz. At higher frequencies the model must be able to simulate a distributed structure. Several attempts were made in the past. So Ziel and Ero [2] represented the planar FET by an active, inhomogeneous, lossy and capacitively loaded line. However, series resistance and field dependent mobility were ignored. Ladbrook [3] describes the channel as a quasi-TEM microstrip line with attenuation $\alpha=0$. Kurvás [4] demonstrates in his model for an unilateral transistor the influence of the gate losses. The influence of the mutual coupling of the gate, drain and source electrodes was first regarded by Ren and Hartnagel [5].

2. FET MODEL WITH DISTRIBUTED ELEMENTS AND WAVE PROPAGATION

In this paper all elements and effects of a FET

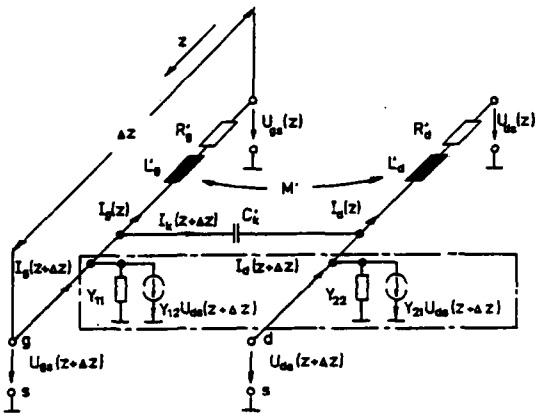


FIGURE 1

Equivalent circuit for a lossy, distributed gate, drain source structure of length Δz

with distributed representation are considered as shown in a section of length Δz in figure 1. The model includes series losses, inductive and capacitive mutual coupling between gate, drain and source as well as field dependent mobility.The transistor equivalent elements are represented by the [Y]-matrix The solution of the problem is described in the following steps:

- Set up of the four differential equations for currents and voltages from the equivalent circuit of figure 1 with mutually coupled electrodes.
- Derivation of the wave equations for the quasi stationary solutions of $\underline{U}_{gs}$ and $\underline{U}_{ds}$.

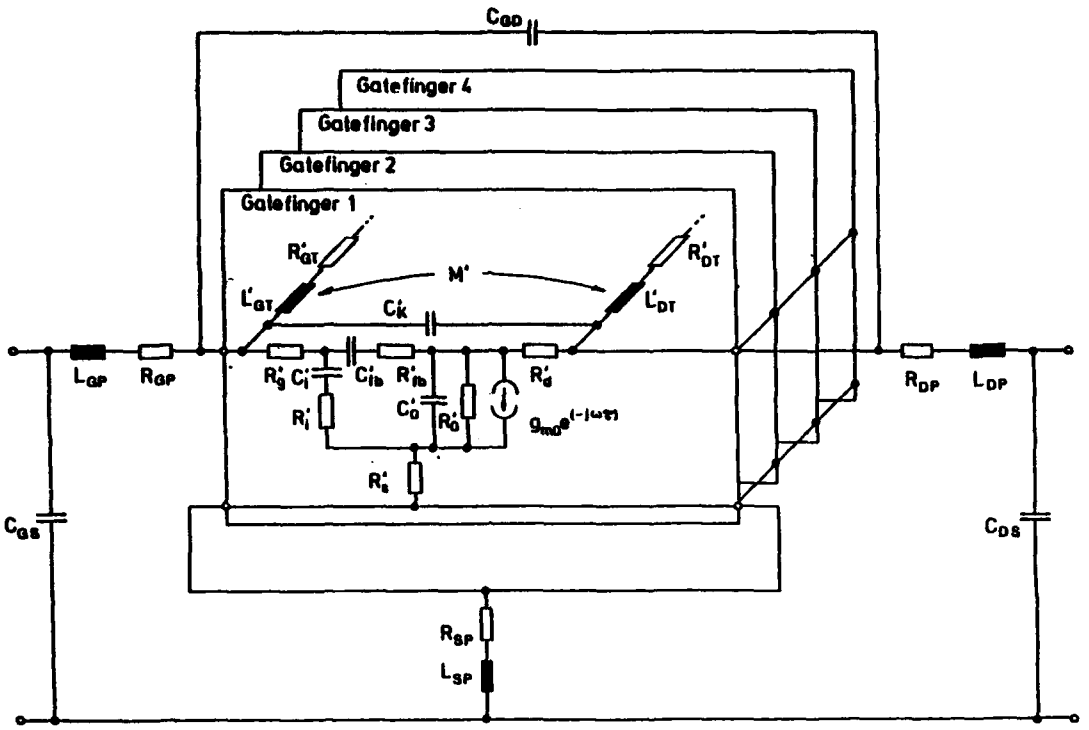


FIGURE 2

Schematic model of a transistor chip with distributed elements

- Decoupling of the above wave equations leads to ordinary, homogeneous differential equations of the fourth order with complex coefficients.

The solution of the eigenvalue leads to the characteristic equations, from which the complex propagation constants γ_{fast} and γ_{slow} for the two waves are derived. The corresponding linear combinations result in the general solution of the problem for $\underline{U}_{gs}(z)$ and $\underline{U}_{ds}(z)$, from which in turn the currents may be derived. With voltages and currents can be computed the characteristic impedances of the resulting four-port and the four-port impedance matrix .

To apply this general solutions to a standard GaAs MESFET with two gate pads (e.g. NE 04500) and a symmetrical connection, four open-ended gate lines have to be connected in parallel. Each of them is actively and passively coupled with drain and source as demonstrated in figure 2.

3. RESULTS

Several types of GaAs MESFETs where used for verification of the above described model. The model contains the basic equivalent circuit elements for the inner transistor and the surrounding structure in the [Y]-matrix. These equivalent circuit elements are determined by the measurement of the bonded chip S-parameters in the frequency range of 1 GHz to 18 GHz and a fitting routine with an evolution optimization. The model for this fitting routine is a small signal representation for the bonded transistor with a maximum of 23 equivalent elements. As far as applicable the equivalent elements are used in the [Y]-matrix. The results of the following diagrams were derived for a NE 04500. They all rely on measured S-parameters up to 18 GHz. To illustrate how active and passive coupling of the lossy gate, drain and source lines with finite length influences the tran-

sistor, measured S–parameters are extrapolated up to 40 GHz with two models:

- first the small signal equivalent circuit model is used for parameter fitting with concentrated elements.
- second this theory with distributed elements and mutual active and passive coupling between gate, drain and source.

The inner transistor model does not differ significantly in both theories. Figure 3 shows the results. A major influence can be recognized for input and output reflection coefficients. As expected the distributed theory results in a higher frequency dependence. Due to the coupling the isolation gets worse. Viewed overall it can be argued that above 30 GHz for this type of FET a distributed model has to be used.

On the same basic as mentioned before, the distributed model is used to calculate the frequency dependence of the propagation constants γ_{fast} and

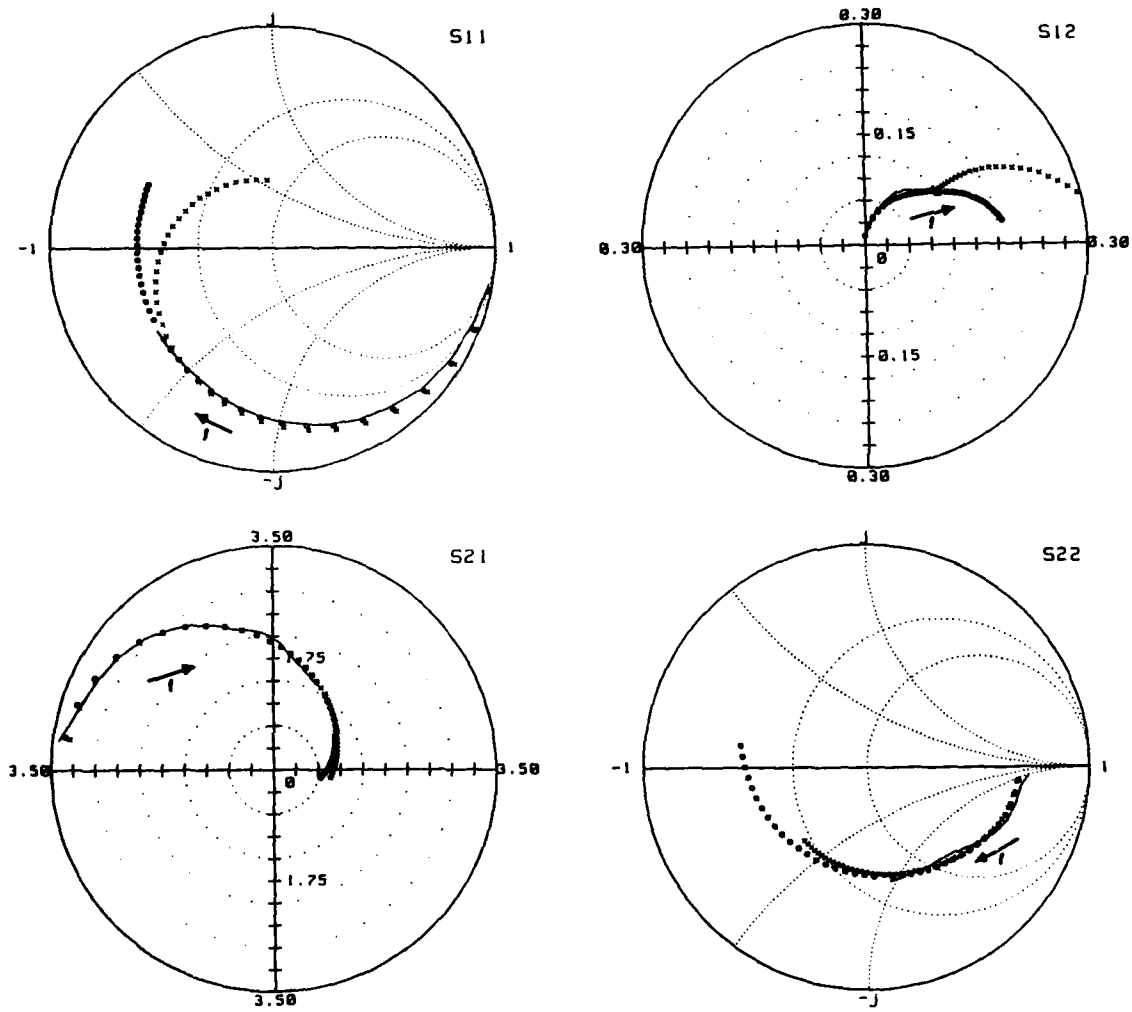


FIGURE 3

Comparison of modelled and measured S-parameters
------- measured 1 - 18 GHz
o o o concentrated elements for modelling 1 - 40 GHz
x x x this theory with distributed elements

γ_{slow}.

Finally the normalized gate and drain voltages are computed for frequencies up to 200 GHz. The same equivalent circuit elements are used as before, but the gate width (W_G) is set from 200 µm to 800 µm (W_{Gu} = 1/4 W_G = 200 µm). The results are shown in figure 4. Significant effects can be recognized at around 150 GHz, far beyond the use of these transistors today. There the gate voltage is dropping, while the drain voltage increases, a typical resonant effect.

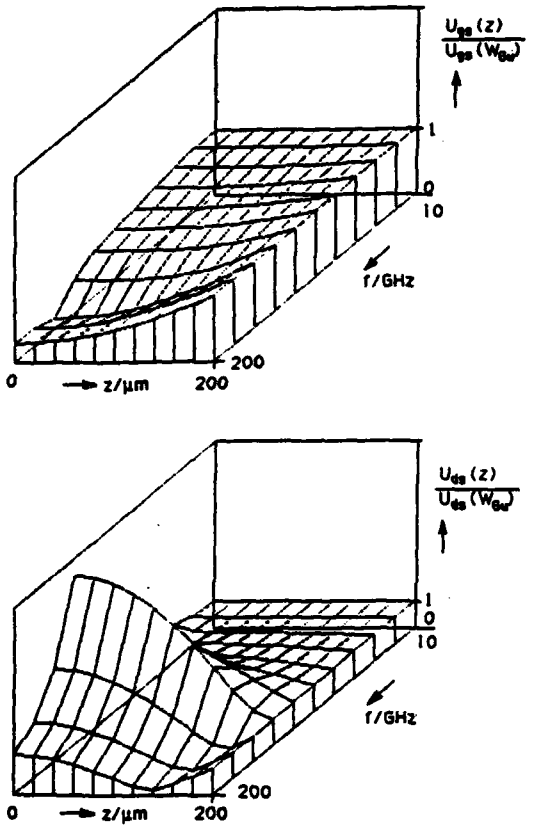


FIGURE 4

Normalized gate and drain voltage versus frequency and gate position

4. CONCLUSION

The comparison of several models for GaAs MESFETs, HEMTs and similar structures has shown, that their validity to higher frequencies may be limited. This is true above 20 GHz to 30 GHz, when mutual coupling and the distributed character of the gate-, drain- and source structure is neglected. At far higher frequencies resonance may appear in the channel. Further studies should involve the inhomogeneous structures and higher order propagation modes.

REFERENCES:

[1] Feldle,H.P.: Klein- und Großsignal- CAD- Modelle von GaAs-MESFETs mit verteilten Elementen, PD-Thesis, Jan. 19. 1987, University of Karlsruhe, Karlsruhe, Germany

[2] Van der Ziel, A.; Ero,J,W.: Small-signal,high frequency theory of field- effect transistors, IEEE Trans. on ED-11, April 1964, pp 128-135

[3] Ladbrooke, P.H.: Some Effects of Wave Propagation of a Microwave MESFET, Electronics Letters, Vol. 14, No. 1, Jan. 1978, pp 21-22

[4] Kuvás, R. L.: Equivalent circuit model of a FET, including distributed gate effects, IEEE Trans. on ED-27, No. 6, June 1980, pp 1193-1195

[5] Ren Y. A.,Hartnagel, H. L.: Wave propagation studies on MESFET electrodes, Int. J. Electronics, Vol. 51,No. 5, 1981, pp 663-668

Session D2.1

Gettering I

Chairman: W. Orr-Arienzo

Tuesday, September 15, 1987

INTRINSIC GETTERING : SENSE OR NONSENSE ?

J. Vanhellemont and C. Claeys

Interuniversity Micro-Electronics Center (IMEC),
Kapeldreef 75, B-3030 Leuven, Belgium

The implementation of intrinsic gettering in integrated circuit processing is discussed in view of new insights in the parameters influencing precipitate formation in silicon and in the dependence of the yield stress on point defect concentrations.

1. INTRODUCTION

During integrated circuit processing metallic contamination is inevitably introduced onto the silicon surface and in the near surface layer. Most of these unwanted impurities are fast diffusers but have a solubility that decreases rapidly with the temperature leading to the formation of a large number of small precipitates during cooling after high temperature indiffusion. The presence of these precipitates, which are very efficient lifetime killers, has a detrimental influence on the electrical performance of the integrated circuits. Several techniques have been developed to getter these metallic impurities in areas where their presence has no influence on the operation of the devices. An extensive review on the interaction of extrinsic and intrinsic point defects and on their influence on dislocation phenomena in silicon can be found in reference [1]. We will limit ourselves here to a brief discussion of the problems related to the implementation of intrinsic gettering in complete device processing.

2. THE GETTERING MECHANISM

Recently Ourmazd [2] showed that most gettering techniques although at first sight of completely different nature, can be understood by assuming the same gettering mechanism consisting of two basic steps. During the first step a concentration pulse of silicon self-interstitials is created which causes a dissolution of the metallic precipitates present in the active areas. These mobilized metal atoms then diffuse rapidly towards an intentionally introduced strained area where they reprecipitate mostly in the form of some type of silicide phase. Based on this idea a more quantitative description of gettering was recently developed by the present authors [3]. This novel approach uses and extends a general expression for the critical radius for coprecipitation of silicon and one type of extrinsic point defect (e.g. interstitial oxygen) [4] to include also the coprecipitation of two (or more) types of extrinsic point defects (e.g. oxygen and carbon) with silicon. For spherical precipitates P in a matrix M, formed by the coprecipitation in a compound $M_y(P_1)_{z1}(P_2)_{z2}\cdots(P_n)_{zn}$ of n "precipitants" P_i ($i = 1,\ldots,n$) present in a concentration C_{Pi}, this leads to the following general expression for the critical radius r_c

$$r_c = \frac{2\sigma}{\frac{EkT}{\Omega_p} \ln \Pi \left(\frac{C_V}{C_V^*}\right)^{\beta} \left(\frac{C_I}{C_I^*}\right)^{\gamma} - 6\mu\delta\varepsilon}$$

$$\Pi = \left(\frac{C_{P_1}}{C_{P_1}^*}\right)^{z_1} \left(\frac{C_{P_2}}{C_{P_2}^*}\right)^{z_2} \ldots \left(\frac{C_{P_n}}{C_{P_n}^*}\right)^{z_n} \qquad (1)$$

with σ the interface energy per unit area, Ω_p the volume per precipitate molecule, γ and β the number of self-interstitials and vacancies absorbed from the matrix per precipitating molecule, μ the shear modulus, δ the linear misfit, ε the "constrained strain", and $E = (1 - \varepsilon)^{-3}$.

By considering the sign of the volume change $\Delta\Omega = \Omega_p - y\Omega_M$ for the different possible precipitate phases, one can predict which phase will dissolve by an intrinsic point defect pulse. The character of the intentionally strained area, i.e. tensile or compressive stresses, then determines the phase in which the mobilized contaminants will reprecipitate. This strained getter region can be located at the surface (frontside gettering), in the bulk silicon (intrinsic gettering) or at the back of the wafer (backside gettering).

3. INTRINSIC GETTERING

Intrinsic gettering is based on the simultaneous generation of a self-interstitial pulse and of a highly defective and strained layer in the bulk of the silicon wafer typically some 20 μm below the surface, by a controlled precipitation of interstitial oxygen. For this purpose high oxygen content wafers are used, also with the hope of keeping some of the superior mechanical strength of this material during processing. The defect free surface layer (the denuded zone) is created by a high temperature oxygen outdiffusion step to create an oxygen lean layer where during the subsequent lower temperature steps no oxygen precipitation and thus no bulk defect generation will occur. Mostly this high temperature oxygen outdiffusion step is followed by a low temperature precipitate nucleation step which will determine the oxygen precipitation kinetics and thus the gettering efficiency during processing. By fine tuning the (low = soaking)/high/low temperature sequence one can obtain a wide range of bulk defect types and densities.

3.1 YIELD STRESS EVOLUTION DURING SELF-INTERSTITIAL INJECTION

A problem that has been somewhat overlooked is the fact that although a denuded zone free from extended bulk defects is obtained, the high mobility and long lifetime of the self-interstitials that are massively generated in the bulk will also increase the self-interstitial concentration in the active areas of the devices. This increase of the self-interstitial concentration C_I above the thermal equilibrium value C^*_I will increase the chemical (and thus the total climb) force on a dislocation nucleus. In defect free substrates under constant external stress, dislocation nucleation is governed by the internal climb force

$$F_{ni} = \frac{b_e kT}{\phi l^3} \ln \frac{C_I}{C_I^*} \qquad (2)$$

with b_e the edge component of the Burgers vector $\mathbf{b}$, l the interatomic spacing between the (111) glide planes, and ϕ a

constant depending on the crystal structure (for Si, $\phi \approx 1$).

The general yield formula in case of an inert anneal of oxygen rich CZ silicon is given by [5,6]

$$F^{y}_{ne}(T) = F^{y0}_{ne}(T) - \frac{bk}{\phi l^3} T \ln \frac{C_I(T)}{C^0_I}$$

$$F^{y0}_{ne}(T) = F^{y}_{nFZ}(T) - \frac{bk}{\phi l^3} T \ln \frac{C^0_I}{C^*_I(T)} \quad (3)$$

with F^{y}_{ne} the external climb force required for dislocation generation (= yielding), F^{y}_{nFZ} the external climb force for yielding in perfect FZ material, F^{y0}_{ne} the external climb force for yielding at time t = 0, C^0_I the concentration of self-interstitials at the beginning of the thermal anneal.
Analytical expressions for C_I can be obtained for inert anneals of oxygen-rich Czochralski silicon [6].
This influence of the amount of precipitated oxygen ΔC on the yield stress is illustrated in figure 1.

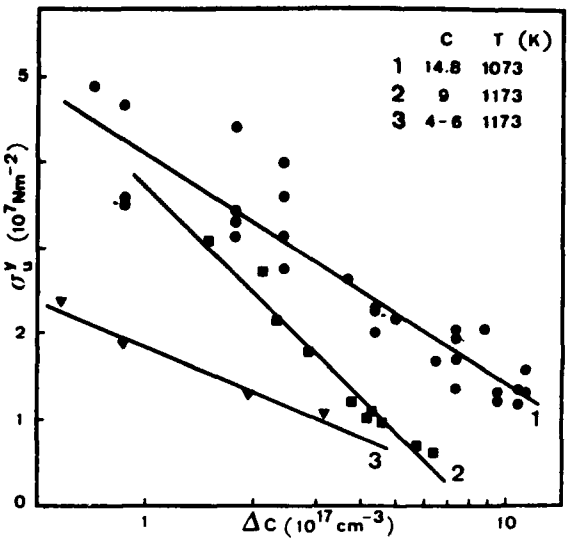


Figure 1

Observed upper yield stress σ^{y}_{u} as a function of ΔC and of the starting interstitial oxygen content C [6].

Intrinsic gettering thus tries to reconcile two conflicting mechanisms resulting from a silicon interstitial concentration pulse: a beneficial one which consists of the dissolution and gettering of metallic precipitates and a detrimental one by the lowering of the yield stress by the increased self-interstitial concentration. For each type of technology and processing conditions there will thus be an optimum of precipitated oxygen ΔC_{Opt} resulting in the highest device yield. This is schematically illustrated in figure 2a which is based on the work of Jastrzebski et al [7]. For $\Delta C < \Delta C_{Opt}$, the gettering action is incomplete while for $\Delta C > \Delta C_{Opt}$ the increased density of processing or warpage induced substrate defects (figure 2b) lowers the overall device yield.

3.2 INTRINSIC GETTERING STRATEGIES

To dissolve the metallic impurities by an increase of the self-interstitial concentration two strategies can be followed : either one opts for a rapid oxygen precipitation at the start of, or immediately after, a "dirty" processing step, or one chooses a more gradual precipitation throughout the complete processing sequence. It is clear that the first option, which one should choose when metallic precipitates are already present, increases strongly the possibility of homogeneous defect nucleation which can result in a lower device yield than compared with the gradual one. This was indeed observed experimentally for bipolar processing as illustrated in figure 2a.

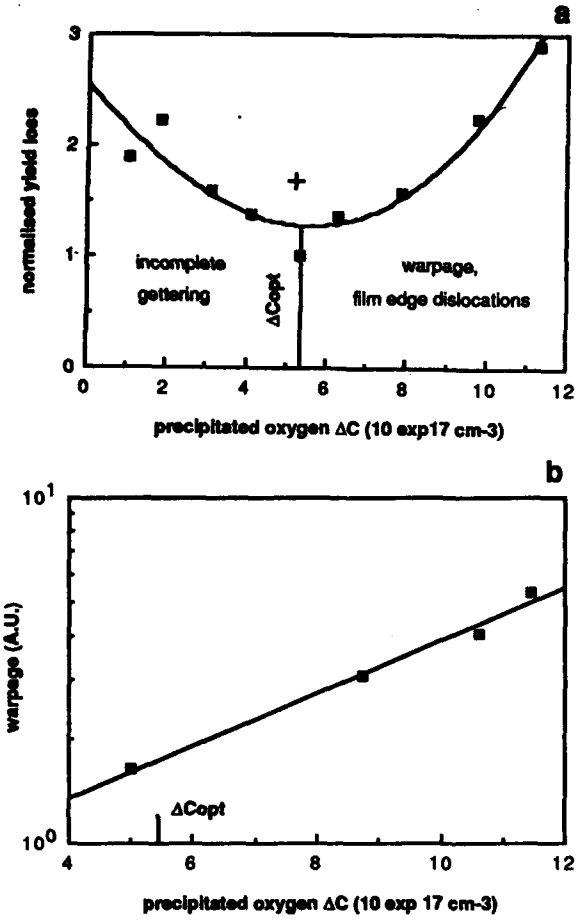


Figure 2

a) Normalised yield loss as a function of the amount of precipitated oxygen during a bipolar process. The cross corresponds with rapid precipitation of oxygen, the other measurements are for gradual precipitation. b) Warpage during bipolar processing as a function of ΔC [7].

The gradual approach should be used to prevent any formation of impurity precipitates during processing of clean substrates. By keeping a (not too) high supersaturation of self-interstitials, the metallic impurities remain in their mobile interstitial status and will diffuse rapidly to the SiO_x precipitates and the associated lattice defects.

4. CONCLUSION

The present paper illustrates that the implementation of intrinsic gettering in complex processing sequences is not that straightforward. The decrease of the yield stress by the intrinsic point fluxes that are generated during oxygen precipitation may jeopardise what is gained by the dissolution and removal of contaminants from the electrically active areas. A gradual oxygen precipitation approach with an adaptation of the amount of precipitated oxygen to the level of metallic contamination seems to be the most promising one.

REFERENCES

[1] Claeys, C. and Vanhellemont, J., Proc. 2nd Int. Autumn Meeting GADEST 1987, Garzau-Berlin(DDR), October 12-17, 1987.

[2] Ourmazd, A., Mat. Res. Soc. Symp. Proc. **59** (1986) 331 .

[3] Vanhellemont, J. and Claeys, C., to be submitted to J. Appl. Phys..

[4] Vanhellemont, J. and Claeys, C., accepted for publication in J.Appl.Phys..

[5] Vanhellemont, J., Claeys, C. and Van Landuyt, J., in "Defects in Semiconductors", ed. von Bardeleben, H.J., Mat. Sc. Forum Vol. 10-12 (1986) 757 .

[6] Vanhellemont,J. and Claeys,C., accepted for publication in J. Electrochem. Soc..

[7] Jastrzebski, L., Soydan, R., McGinn, J., Kleppinger, R., Blumenfeld, M., Gillespie, G., Armour, N., Goldsmith, B., Henry, W. and Vecrumba, S., J. Electrochem. Soc. **134** (1987) 1018 .

LIFETIME ENGINEERING BY OXYGEN PRECIPITATION IN SILICON*

M. L. Polignano and G. F. Cerofolini

SGS Microelettronica, 20041 Agrate MI, Italy

H. Bender and C. Claeys

IMEC, Kapeldreef 75, 3030 Heverlee, Belgium

J. Reffle

Wacker Chemitronic, 8263 Burghausen, West Germany

This work studies the preannealing steps at high (HI) and low (LO) temperatures for internal gettering in medium oxygen-content silicon slices [$(7 \pm 1.5) \cdot 10^{17}$ oxygen atoms/cm^3]. By using silicon p-n junctions as test vehicles, we get the following conclusions: 1) LO–HI pretreatments produce an increase of defects with respect to non-preannealed wafers; such an increase is not observed in HI–LO pretreatments; 2) for HI–LO preanneals an optimum high temperature treatment can be found, which gives lower amount of defects with respect to non-preannealed wafers, reproducibly from lot to lot; 3) in HI–LO preannealed wafers the recombination lifetime increases with increasing the oxygen diffusion length of the high temperature pretreatment; the electron diffusion length in the bulk decreases with increasing initial interstitial oxygen concentration; 4) oxygen precipitates are not effective getter sites for metal impurities.

1. INTRODUCTION

In this work we study: 1) pre-annealing steps for the formation of a high quality denuded zone (DZ) in medium oxygen content silicon; 2) the defectiveness of 'optimum' DZ compared to the one of non-preannealed wafers; 3) the influence of initial interstitial oxygen concentration [O_i] on the electrical performance of the test devices; 4) the gettering effectiveness by oxygen precipitates compared to the one by doping and segregation annealing.

The test devices are two n^+-p junctions, one with an area $6.25 \cdot 10^{-3}$cm^2 and a perimeter 1 cm ('area diode') and the other with area $1.25 \cdot 10^{-3}$cm^2 and perimeter 62.4 cm ('perimeter diode'). The diode process, listed in table 1, includes an external gettering by phosphorus predeposition and segregation annealing at 800°C for 1 h.

The diodes have been tested in reverse bias condition to determine the layer defectiveness: the most frequent value of leakage current is assumed as the typical value of the process and a diode is defined to be defective when its reverse current exceeds by more than one order of magnitude the typical value. Both forward and reverse characteristics of the typical diode have been obtained and elaborated to get generation-recombination lifetime in the denuded zone and the electron effective diffusion length.

Table 1: n^+-p diode process.

- First oxidation and Si_3N_4 deposition
- Mask (defines active zone) and field ion implant
- Field oxidation, 920°C, $x_{ox} = 7000$Å
- Si_3N_4 etching, first oxide etching
- Gate oxidation, gate oxide annealing, oxide etching
- As ion implant, $6 \cdot 10^{15}$ cm^{-2}, 150 keV
- Radiation damage anneal, 550°C, N_2, 2 h
- Oxidation, 800°C, 50 min, $x_{ox} = 300$Å
- $SiO_2 : P_2O_5$ deposition and densification
- Mask (defines contacts), oxide etching, back oxide etching
- P predeposition, $POCl_3$, 920°C, $V/I = 4\,\Omega$
- Reflow anneal, 1050°C, N_2, 5 min
- Segregation anneal, 800°C, N_2, 1 h
- Al:Si sputtering, mask, Al:Si etching and alloy

The final junction depth is about 0.35μm.

2. DZ FORMATION

The slices were p-type, Czochralski grown, (100) oriented, 10 cm diameter, 16 – 24 Ω cm resistivity, and [O_i] was in the range $7 \pm 1.5 \cdot 10^{17}$cm^{-3} (measured by infrared absorption spectrometry, new ASTM units); carbon concentration was below or around the detection limit (10^{16}cm^{-3}).

The preliminar experiment consisted in comparing LO–HI preannealed diodes with HI–LO ones [1,2]. The preliminar choice was: LO: 730°C, 8 h N_2; HI: 1050°C, 2 h, 4 h, 8 h or 1100°C, 55 min, 1 h 45 min, 3 h 35 min, N_2 + 5% O_2 or

*This work was partially supported by the European Economic Community under the ESPRIT program "Substrates for CMOS VLSI".

100%O_2. The durations of the high temperature treatments were so chosen that the oxygen diffusion lengths at 1050°C and 1100°C were equal to one another [3] (2 h at 1050°C ↔ 55 min at 1100°C, etc). LO–HI pretreatments produce an increase of both area and perimeter defect density (δ_A and δ_P; see table 2).

Table 2: Defect density of preannealed wafers.

	δ_A/cm^{-2}	δ_P/cm^{-1}
LO–HI	1.7	$8 \cdot 10^{-3}$
HI–LO	0.6	$4 \cdot 10^{-3}$
not preannealed	0.9	$6 \cdot 10^{-3}$

Area and perimeter defect density of preannealed and non-preannealed wafers. About 500 diodes per group were tested.

The second step concerned the analysis of the effect of the duration of the HI treatment [(2 – 8)h at 1050°C and 55 min – 6 h at 1100°C] in HI–LO sequences. It has been found that this preannealing has a rather small effect on area defect density, but a sensitive effect on perimeter defect density (see table 3): in wafers preannealed at 1100°C for 6 h (oxygen diffusion length $2\sqrt{D_O t} \simeq 34\mu m$) the perimeter defect density is the lowest and most reproducible; for shorter oxygen diffusion length and in non-preannealed wafers the perimeter defect density is higher and disperded over a wider range. No influence of the annealing enviroment was found.

Table 3: Defect density of HI–LO preannealed wafers.

$2\sqrt{D_O t}/\mu m$	lot-to-lot dispersion	typical value
	δ_P/cm^{-2}	
14	$0.9 - 12 \cdot 10^{-3}$	$3.6 \cdot 10^{-3}$
20	$< 0.4 - 13 \cdot 10^{-3}$	$5.4 \cdot 10^{-3}$
28	$0.9 - 16 \cdot 10^{-3}$	$3.9 \cdot 10^{-3}$
34	$0.4 - 4 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$
not preannealed	$< 0.2 - 55 \cdot 10^{-3}$	$6 \cdot 10^{-3}$

Most electrical defects in LO–HI preannealed wafers are due to a seeming anticipated breakdown, in the reverse bias range 10 – 20 V, the correct breakdown voltage (BV) being 20 – 23 V. The forward characteristic is not affected by this defect. This behaviour is sometimes observed also in non-preannealed wafers, but usually in perimeter diodes. HI–LO pretreatments reduce the number of such defects, provided that the duration of the HI annealing is long enough, and this results further in a reduction of perimeter defect density (table 3).

The seeming anticipated breakdown is much less frequent in epitaxial substrates than in standard, non-preannealed substrates: for epitaxial substrates $\delta_P = 3 \cdot 10^{-4}$–$4 \cdot 10^{-3} cm^{-1}$, with an average value of $2 \cdot 10^{-3} cm^{-1}$, comparable with the value obtained in HI–LO preannealed wafers by the longest HI pretreatment. These facts suggest that the seeming anticipated breakdown is related to the presence of oxygen in a layer close to the surface.

Structurally, the layer has been observed both by scanning electron microscope (SEM) and by transmission electron microscopy (TEM); when precipitates are large enough, these techniques give the same value of denuded zone (DZ) thickness. Typical defects observed by TEM in the bulk of LO–HI preannealed wafers are: plate-like amorphous SiO_x precipitates with prismatic punching, small precipitates (500 – 700 Å) with density $10^{12} cm^{-3}$, large dislocation loops, and stacking faults with diameter 1–10μm (fig. 1). Also in HI–LO preannealed wafers oxide precipitates with prismatic punching and irregularly shaped dislocation loops are revealed, but bulk defects are much less dense. The density of bulk defects is sometimes very low, so that it cannot be estimated by TEM inspection; a bulk defect density of the order of $10^9 cm^{-3}$ has been found for initial interstitial oxygen concentration of $8 \cdot 10^{17} cm^{-3}$ (close to the upper limit of the range in study).

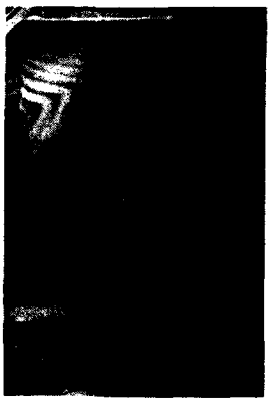


FIGURE 1

Typical defects in the bulk of a LO–HI preannealed wafer.

The minority carrier effective diffusion length in the neutral bulk region, L_{eff}, has been obtained from the diffusion component of the forward current density of the diode; both for LO–HI and for HI–LO pretreatments L_{eff} increases with the oxygen diffusion length of the HI pretreatment

(table 4). For LO–HI pretreatments L_{eff} is comparable to the denuded zone thickness L_{DZ}, while for HI–LO pretreatments L_{eff} is larger by one order of magnitude. By assuming $2\sqrt{D_O t} = L_{DZ}$, the electron diffusion length in the defective bulk zone L_b can be calculated; L_b is in the range 1–10μm for LO–HI pretreatments and 50–200μm for HI–LO pretreatments.

Table 4: Denuded zone thickness.

	LO–HI		HI–LO	
$2\sqrt{D_O t}$	L_{DZ}	L_{eff}	L_{DZ}	L_{eff}
14	17.5	20 – 23	16	90 ± 30
20	22	36	20	140 ± 30
28	29	27 – 33	30	150 ± 50
34			39	170 ± 70

Denuded zone thickness measured by cleavage, Secco etching and SEM inspection is compared to the oxygen diffusion length for the high temperature pretreatment and to the effective electron diffusion length. All lengths are expressed in μm. The typical dispersion in L_{DZ} is of the order of a few microns.

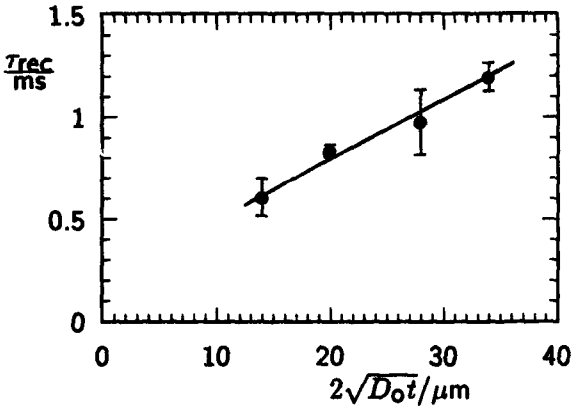


FIGURE 2

Recombination lifetime in HI–LO preannealed wafers vs. the oxygen diffusion length for the high temperature preannealing.

The recombination lifetime in the space charge region τ_{rec} has been obtained from the recombination component of the forward area current density. For HI–LO preannealed wafers τ_{rec} increases when the oxygen diffusion length increases (fig. 2); the typical value of non-preanneled wafers (about 1 ms) is reached after 6 h at 1100°C. For LO–HI preanneals the recombination lifetime is in the range $2–3 \cdot 10^{-5}$s.

The previous results are an average over all oxygen contents in the specified range. The subsequent analysis concerned the role of oxygen concentration. Slices with measured oxygen content (in the same range as above) were preannealed according to the 'best' pretreatment: a) 1100°C, 6 h, N_2 + 5%O_2; b) 730°C, 8 h, N_2; c) 1000°C, 5 h, 100% dry O_2. When $[O_i]$ varies in the specified range, no significant variations are observed in defect density of preannealed wafers and in the percent variation of interstitial oxygen concentration $\Delta[O_i]$ after the preannealing steps ($\Delta[O_i] \leq$ 5–10%). Also the recombination lifetime in the space charge region is independent of $[O_i]$ in the considered range, varying from $6 \cdot 10^{-4}$s to $8 \cdot 10^{-4}$s.

However, the electron diffusion length in the bulk is strongly influenced by $[O_i]$ (fig. 3): for preannealed wafers processed according to table 1 L_b decreases with increasing oxygen concentration for $[O_i] > 7 \cdot 10^{17}$cm^{-3}, starting from the typical value of non-preannealed wafers.

3. GETTERING BY OXYGEN PRECIPITATES

In order to evaluate the effect of the segregation annealing (SA) [4,5] on the behaviour of preannealed wafers, we selected preannealed wafers with the same values of $[O_i]$ and $\Delta[O_i]$; some of them did not undergo the segregation anneal-

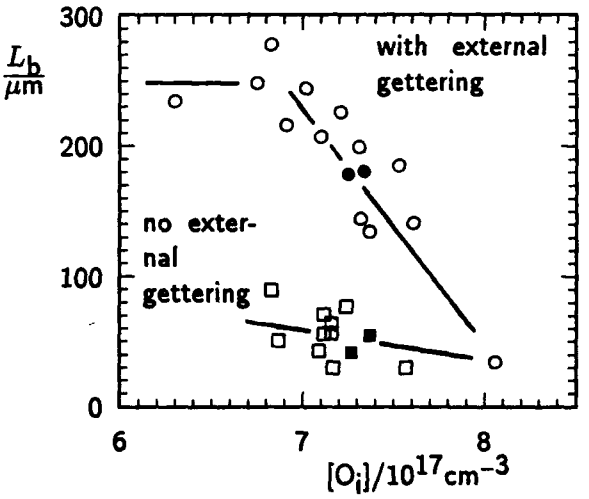


FIGURE 3

Electron diffusion length L_b in the bulk of wafers preannealed according to the optimised preannealing sequence. vs. initial $[O_i]$. o and •: process shown in table 1; □ and ■: the same, except that the final external gettering step (phosphorus predeposition and segregation annealing) was not carried out. Samples • and ■ had the same value of $[O_i]$ and of $\Delta[O_i]$ ($\Delta[O_i] = 0.5 \cdot 10^{17}$cm^{-3}). L_b was calculated assuming a 50μm denuded zone (obtained from TEM inspection).

ing step. In these wafers L_b is lower than a half the value in wafers which underwent segregation annealing (fig. 3). Similar results are obtained if no phosphorus predeposition but only a segregation annealing is performed or if no gettering by doping and segregation annealing is performed. These results show that *gettering by phosphorus doping and segregation annealing is more effective than gettering by oxygen precipitates*, at least for initial interstitial oxygen concentration in the considered range.

Table 5: External gettering in HI–LO preannealed wafers.

preanneal	final gettering steps	$\frac{\delta_A}{cm^{-2}}$	$\frac{\delta_P}{cm^{-1}}$
yes	P predep + SA	0.6	$1.7 \cdot 10^{-3}$
yes	no P predep, only SA	15	$3 \cdot 10^{-2}$
yes	no final gettering	15	$2 \cdot 10^{-2}$
no	P predep + SA	0.9	$6 \cdot 10^{-3}$
no	no final gettering	14	$> 5.5 \cdot 10^{-2}$

Area and perimeter defect density in HI–LO preannealed and non-preannealed wafers for different choices of the final gettering steps. The optimized preannealing sequence has been used.

The results in fig. 3 concern only the behaviour of good diodes obtained by different gettering processes. Besides, in absence of phosphorus predeposition the defect densities are very high both in preannealed and in non-preannealed wafers (table 5); only a small reduction of δ_P is obtained by the preannealing sequence. When no gettering by phosphorus predeposition and segregation annealing is performed, both in preannealed and in non-preannealed wafers defective diodes often show a 'soft' reverse characteristic (attributed to metal contamination [6]), which is very seldom observed in presence of phosphorus gettering. These facts show that *gettering by oxygen precipitates alone is not able to prevent from electrical defects.*

As LO–HI pretreatments have been found to produce a higher density of bulk defects than HI–LO ones, more gettering efficiency could be expected from the defective bulk zone created by a LO–HI preanneal; in fact, we observed decorated bulk defects by TEM inspection in absence of phosphorus gettering. Despite this, table 6, which compares the recombination lifetime in the space charge region obtained in LO–HI preannealed and non-preannealed wafers with and without gettering by phosphorus predeposition and segregation annealing, shows that also a heavily defective bulk zone, as the one created by a LO–HI preanneal, has a very poor gettering efficiency. In fact, in absence of phosphorus gettering LO–HI preannealed wafers have low recombination lifetime; these wafers give a leakage current of $1\mu A/cm^2$, which is unacceptably high for several microelectronic applications.

Table 6. External gettering in LO–HI preannealed wafers.

	τ_{rec}/s	
	LO–HI	no preanneal
P predep + SA	$2\text{–}3 \cdot 10^{-5}$	$1\text{–}3 \cdot 10^{-3}$
no final gettering	$2 \cdot 10^{-7}$	$5 \cdot 10^{-7} - 3 \cdot 10^{-5}$

Recombination lifetime in LO–HI preannealed and non-preannealed wafers with and without gettering by phosphorus doping and segregation annealing.

4. CONCLUSIONS

Preannealing steps for the formation of a high quality denuded zone in medium oxygen content silicon have been studied. LO–HI preannealed wafers have been shown to have higher defect density than non-preannealed wafers. The HI pretreatment has been optimised for HI–LO pretreatments; a pretreatment has been identified which gives wafers with lower and more controlled defect density than non-preannealed wafers.

In HI–LO preannealed wafers the electron diffusion length in the bulk decreases with increasing initial interstistial oxygen concentration; the recombination lifetime in the denuded zone increases with increasing oxygen diffusion length.

The gettering efficiency of the defective bulk zone has been tested; it has been found that oxygen precipitates are not effective gettering sites for metal impurities, and gettering by doping and segregation annealing is still necessary to control the electrical performances of devices. Therefore, the main result of a correct preannealing sequence is to prevent from oxygen precipitation in the surface layer.

REFERENCES

1. T. Y. Tan, E. E. Gardner and W. K. Tice, Appl. Phys. Lett. **30** (1977) 175
2. E. M. Murray, J. Appl. Phys. **55** (1984) 536
3. F. Shimura and H. Tsuya, J. Electrochem. Soc. **129** (1982) 1062
4. L. Baldi, G. F. Cerofolini, G. Ferla and G. Frigerio, Phys. Stat. Sol. (a) **48** (1978) 523
5. L. Baldi, G. F. Cerofolini and G. Ferla, J. Electrochem. Soc. **127** (1980) 164
6. A. Goetzberger and W. Shockley, J. Appl. Phys. **31** (1960) 1821

A MODEL FOR OXYGEN PHASE TRANSITION KINETICS IN CZ GROWN SILICON AND ITS APPLICATIONS TO IC PROCESSES

M.Pagani

Dynamit Nobel Silicon SpA, 28100 Novara, Italy

W.Huber

Dynamit Nobel Silicon Technology Center, Sunnyvale, CA 94086

A model for oxygen precipitation in silicon at high temperatures has been developed and tested succesfully on different low-high treatment. The model was then extended to multistep treatments and tested on CMOS/NMOS processes simulations ; trends are correctly predicted.

1. INTRODUCTION

VLSI device manufacturing requires tight control of physical features of silicon wafers,that becomes very important for oxygen for its impact on device attributes. Oxygen content in Si can be controlled anywhere between $5*10^{17}$-10^{18} at cm^{-3}. Such values exceed solubility at all process temperatures and oxygen precipitates forming a Si-O phase[1,2] during device processing. Precipitates have a negative effect in the active device zone where are linked to generation/recombination centers[3,4],while they act as gettering centers for impurities in the bulk.

Modeling of oxygen precipitation (OP) is then of great interest because it permits to tailor the wafer properties for maximum device yield in a given process without costly experiments.

In this work OP has been studied for different initial conditions and thermal treatments (TT). The effect of a 2-step TT was firstly studied. A model involving homogeneous nucleation[5] and diffusion limited precipitation[6] has been developed to interpretate the results. In a second time the influence of CMOS/NMOS processes on OP was studied; thermal simulations were performed and the model adapted to multistep TT.

2. EXPERIMENTAL

Samples used in this study were boron doped, (100) oriented polished wafers,125 mm in diameter and 20-60 ohm cm in resistivity. Oxygen content C_o was measured with FTIR spectroscopy according to ASTM F121-83. Samples were chosen to obtain a uniform distribution of C_o between 6 and $8*10^{17}$ at cm^{-3}.

In 2-step TT samples were annealed in N_2 for 0-8 h at 650-850 °C and then for 0-20 h at 1000-1150 °C. CMOS/NMOS simulations reproduced all process thermal steps without lithography,LPCVD. implantations. Oxidations were performed in O_2 or H_2/O_2, other steps in N_2.

Residual oxygen was measured with FTIR spectroscopy. Precipitate density was measured on cleaved cruss sections after 5' Schimmel etch.

3. DESCRIPTION OF THE MODEL

3.1 Precipitation at constant temperature

In order to model OP in MOS processes we will first analize OP in 1 step TT at constant temperature. The following assumption are made:

* OP occurs only if oxygen nuclei (generated during crystal growth or low temperature steps) are present with density N_o.

* OP is caused by nuclei growth via spherical diffusion of oxygen atoms.

* During growth precipitates keep smaller than the mean distance among nuclei.

The precipitated oxygen fraction, defined as $Sp(t)=(C_o-C(t))/(C_o-Cs(T))$, were C(t)=residual concentration at time t, Cs=solubility[7], can then be expressed as

$$(1) \qquad K_o t = H(Sp^{1/3})$$

$$(2) \qquad H(u)=\tfrac{1}{6}\ln\frac{u^2+u+1}{u^2-2u+1} - 3^{\frac{1}{2}}\,\mathrm{artg}\,\frac{2u+1}{3^{\frac{1}{2}}} + 0.9068$$

where K_o is a kinetic constant

$$(3) \qquad K_o = D(T)\, N_o^{2/3}\,(36\ V_o\ (C_o-Cs))^{1/3}$$

where V_o=oxygen atomic volume in the precipitate, D=diffusivity at temperature T. From (1), we see that choosing $K_o t$ as time variable OP kinetics is the same for all TT and initial conditions.

N_o is given by homogeneous nucleation theory,

$$(4) \qquad N_o = Js(t-t_o(1-e^{-t/t_o}))$$

$$(5) \qquad Js = J_o D\, C_o^2\, T^{-\frac{1}{2}}\, e^{-E_o/(T(1-T/Ts)^2)}$$

where t_o=incubation time, Js=nucleation rate, J_o, E_o material constants, Ts=solubility temperature when C_o equals Cs.

3.2 Precipitation during multistep TT

The model is extended to multistep TT with the following assumptions:

* Steps below 850 °C generate nuclei according to (4)-(5), with C(t) instead of C_o, but no OP. Nuclei formed during CZ growth are neglected.

* The effect of temperature ramps is neglected.

* A multistep TT is considered as a single TT; in this way K_o is discontinuous at each step boundary for changes in temperature or N_o.

The residual concentration within each step can be computed from (1) introducing a time shift t_n^* at each step n, defined by

$$(6) \qquad t_n^*=((K_n/K_{n+1})-1)t_n + (K_n/K_{n+1})\, t_{n-1}^*$$

where K_n, t_n kinetic constant and length of step n. t_n^* takes into account that different K_o produce the same OP at different times and that C(t) is continuous at each step boundary.

4. RESULTS AND DISCUSSION

4.1 OP after low-high TT

In fig.1 observed Sp is plotted vs reduced time $K_o t$ using the observed N_o values. All data cluster around the theoretical curve (solid line): this is a first test of the model's validity.

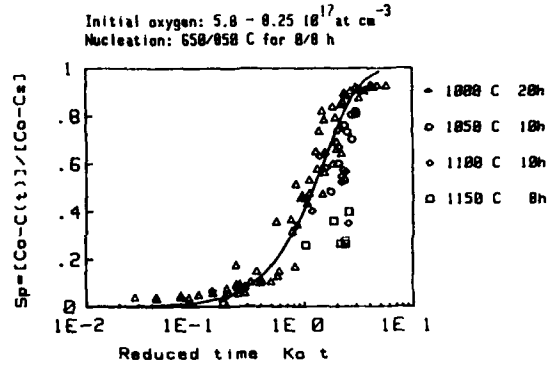


FIGURE 1

Diffusivity calculated from (2)-(3) is in good agreement with published data[8] (fig. 2).

Analysis of oxygen nucleation is beyond the aims of this work, but an extimation of nucleation rate was made to compute nuclei formed during nucleation steps. A best fit of (5) gives $J_o=7.94*10^{-11}$ cm $K^{\frac{1}{2}}$ and E_o=1082 K.

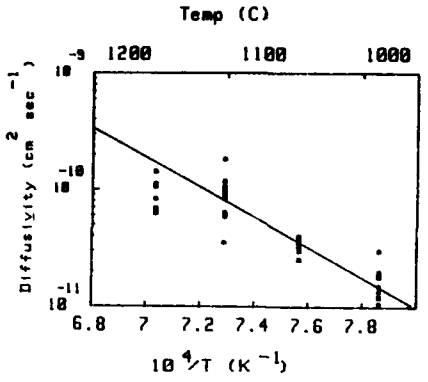


FIGURE 2

4.2 OP after CMOS/NMOS process simulation

A first group of samples was submitted to the CMOS simulation of Table 1. After each step wafers were cooled to room temperature. A second group ("modified process") was treated for 4 h at 700 °C before the simulation,to study the effect of a *prenucleation* on OP;in this case pad oxide was grown at 1100 °C to obtain an acceptable denuded zone depth.

In fig. 3 the mean C(t) is plotted vs process time;after a typical CMOS process OP is 15% while after the modified process OP reaches 60%. The determining step is then the well diffusion.

In fig. 4 precipitated oxygen is plotted vs C_o;solid lines are the model's predictions;OP trend is correctly predicted by the model. Prenucleation reduces OP scattering for a given C_o.

A third group of samples was submitted to NMOS simulation (steps of Table 1 without well diffusion). A modified process was run with a fourth group as for CMOS.

In fig. 5 the mean C(t) is plotted vs process time. After a typical NMOS process OP is negligible,while after the modified process OP reaches 30 %;well diffusion has a strong impact on OP.

In fig. 6 precipitated oxygen is plotted vs C_o. Solid lines are the model's predictions;also in this case OP trend is correctly predicted.

Table 1. Typical CMOS process thermal steps

* Pad oxide, 500 A , 950 °C
* Nitride deposition, 200 A, 780 °C
* Field oxide, 5000 A, 950 °C
* Well diffusion, 1150 °C
* Gate oxide, 300 A, 900 °C
* Polysilicon deposition,5000 A, 650 °C
* Source and drain diffusion, 1000 °C
* PVAPOX deposition, 415 °C
* PVAPOX densification, 900 °C
* Metal forming, 450 °C
* Final passivation deposition, 415 °C

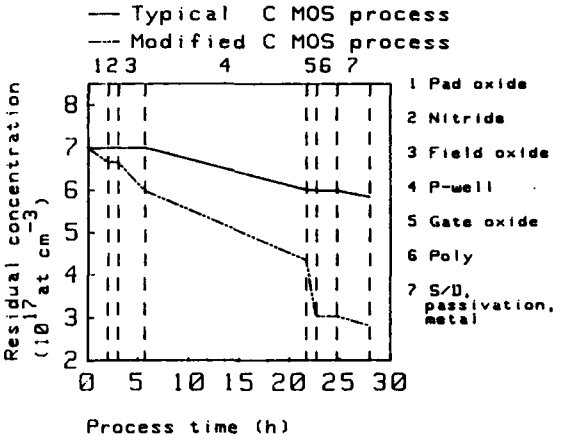


FIGURE 3

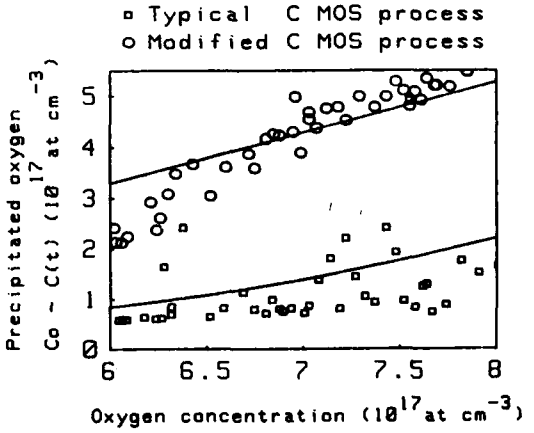


FIGURE 4

Prenucleation doesn't reduce OP scattering,because the shorter NMOS process times don't permit to level OP in samples with different grown in nuclei densities.

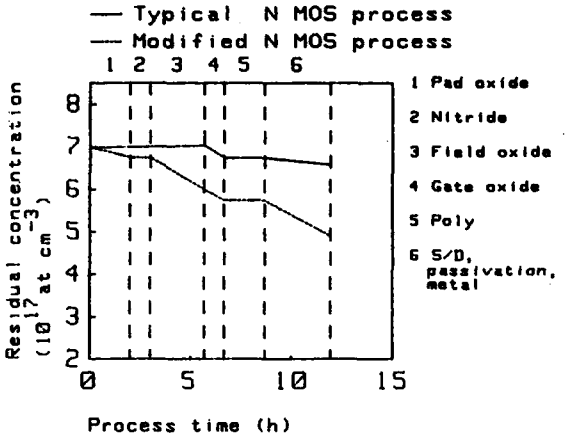


FIGURE 5

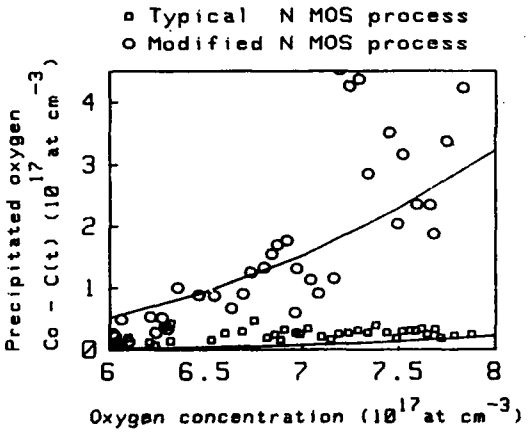


FIGURE 6

5. CONCLUSIONS

A model for OP in silicon has been developed to predict oxygen kinetics during high temperature TT and tested succesfully on2-step TT. The model was then extended to multi-step TT and tested on CMOS/NMOS processes simulations;data are fitted well by the model.

Precipitation scattering can't be explained by the model it's probably due to growth in nuclei; this item has to be implemented.

In any case a good extimation of OP for different process technologies can be made and this permits to design the substrate's characteri - stics in order to maximize OP and intrinsic gettering.

REFERENCES

1. Hu,S.M.,J.Appl.Phys.,51,5945 (1980)

2. Bourret,A.,Thibault-Desseaux,J.,Seidman,D.N., J.Appl.Phys,55,825 (1984)

3. Patel,J.R.,Jackson,K.A.,Reiss,H.,J.Appl.Phys. 48,5279 (1977)

4. Patrick,W.J.,Hu,S.M.,Westdorp,W.A.,J.Appl. Phys.,50,1399 (1979)

5. Inoue,N.,Wada,K. and Osaka,J., Oxygen precipitation in czochralski silicon.Mechanism and application, in: Huff,H.R.,Kriegler,R.J., Takeishi,Y. (eds.), Semiconductor Silicon 1981 (The Electrochemical Society,Pennington, N.J.,1981), p. 282

6. Ham,F.S.,J.Phys.Chem.Solids,6,335 (1958)

7. Craven,R.A., Oxygen precipitation in czochralki silicon, in : Huff,H.R.,Kriegler,R.J.,Takeishi,Y. (eds.), Semiconductor Silicon 1981 (The Electrochemical Society,Pennington,N.J., 1981), p. 254

8. Stavola,M.,Patel,J.R.,Kimerling,L.C.,Freeland,P.E.,Appl.Phys.Lett.,42,73 (1983)

ACKNOWLEDGEMENTS

This work was realized with the partial support of the Italian Consiglio Nazionale delle Ricerche.

INFLUENCE OF INTRINSIC GETTERING ON SILICON RECOMBINATION PROPERTIES AND THEIR RELATION TO DEVICE PERFORMANCE

Martin KITTLER, Hans RICHTER, and Winfried SEIFERT

Academy of Sciences of the GDR
Institute of Semiconductor Physics
W.-Korsing-Str. 2, DDR-1200 Frankfurt (Oder)

1. INTRODUCTION

Intrinsic gettering (IG) procedures in Cz silicon result in the formation of defect depth profiles characterized by a low density of crystal defects at the surface (denuded zone-DZ) and a high defect density in the wafer bulk. Thereby, the bulk defects are important as gettering sites for undesired impurities /1/. The defect profile leads to a corresponding profile of recombination properties (recombination lifetime τ respectively diffusion length L), with low recombination in the DZ and strong recombination in the bulk.

This is illustrated in /2, 3/ showing EBIC micrographs of beveled samples and the corresponding depth profiles of minority-carrier diffusion length L(z). The bulk region is often treated as being unimportant for device operation. This assumpution is, however, not justified when devices are built directly in the substrate /4/.

Accordingly, there is an interest in characterizing bulk recombination in IG silicon and in knowledge about its sources. Studies on this topic performed by the EBIC technique are presented below. Further, the influence bulk recombination may have on device operation will be discussed in chapter 3.

2. RECOMBINATION SOURCES IN THE BULK OF IG SILICON

During IG oxygen - related precipitates (OP) are formed which on its turn lead to generation of secondary defects like volume stacking faults (SF) and dislocation loops, see /1/.

In a paper by Hwang and Schroder /5/ the OP themselves are described as the dominant source of bulk recombination although SF and dislocation loops were also present in the samples investigated by the authors. The general validity of their results is to be called in question from the point of view of our experience with intrinsically gettered silicon samples.

The EBIC method was used by us for investigating the causes of bulk recombination as it allows both to determine diffusion length and to visualize sites of strong recombination. The strength of the defect recombination activity may be thereby evaluated from the defect EBIC contrast /6 - 8/. But, by far not all volume defects reveal themselves by producing EBIC contrast. This is in agreement with results of Inoue et al. /9/ who carried out EBIC investigations of crystal defects in solar cells and found a correlation between local diffusion length and the density of defects showing EBIC contrast.

On the base of EBIC investigations of IG silicon wafers we are going now to rediscuss the question of Hwang and Schroder /5/: What is the dominant source of lifetime degradation, OP or secondary defects like SF or dislocation loops?

It is obvious that defects showing dark EBIC contrasts reduce the avarage diffusion length in the material because dark contrasts indicate additional recombination. So far it is proved that secondary defects with EBIC contrasts are recombination sources. However, other crystal defects not detectable by the EBIC technique and impurities may also produce an enhanced recombination background without producing EBIC contrast (Fig. 1). So there arises the

problem to separate the contributions due to the background from that due to defects having EBIC contrast.

2.1. On the EBIC-technique

Assume recombination-active defects being small compared to the size of the interaction volume of the electron probe with the semiconductor. Then recombination properties and EBIC contrast are related by the formula /8/

$$n\sigma \simeq c \frac{D}{f v_{th}} \qquad (1),$$

$n\sigma$ being the product of number and mean capture cross section of recombination centres acting in the defect, c the EBIC contrast of the defect, D and v_{th} the minority-carrier diffusivity and thermal velocity, respectively, and f a correction factor depending on defect depth, operation conditions of the scanning electron microscope and diffusion length outside the defect. To estimate the different contributions to recombination a simple model will be used here which treats the recombination taking place at the defects visible in EBIC contrast as being equivalent to centres randomly distributed in the material. So, we may write

$$L_B^{-2} = L_E^{-2} + L_M^{-2} \qquad (2)$$

where L_B is the bulk diffusion length measured, L_E describes the contribution of the observed defects with EBIC contrasts and L_M is the diffusion length due to the recombination background, compare Fig. 1. The diffusion length component due to the EBIC contrasts, L_E, depends on the contrast values c and on contrast density and may be roughly approximated by (unpublished result)

$$L_E \simeq l \left(\frac{Rf_{max}}{c_{max}} \right)^{1/2} \qquad (3)$$

with l describing the mean distance between the contrasts on the EBIC micrograph, $R(\mu m) = 0.0171 \times E_o^{1.75}$ (keV) being the electron range /10/ and c_{max} the maximum contrast. Thereby, $Rf_{max} \simeq 0.15$ is valid independent of R for diffusion lengths between the defects $L_M \gtrsim 10\ \mu m$. Fig. 2 shows the bulk diffusion length L_B calculated from (2) and (3) versus mean distance l between EBIC contrasts.

2.2. Experimental results and discussion

There were studied a lot of different Cz silicon samples with the following characteristics: p-type, boron-doped, ϱ = 10 ... 20 ohm.cm, 3" or 4" diameter, (100) or (111) orientation, different oxygen content (5...10 x 10^{17} cm^{-3}), different heat treatments for IG (multistep heat treatments or ramping for nucleation).

The EBIC investigations in the volume of the samples were carried out either on bevels or after sufficient surface removal by etching. Al Schottky contacts were used for charge collection.

l and c_{max} were determined for a beam energy E_o = 30 keV thereby averaging over a sufficiently large area.

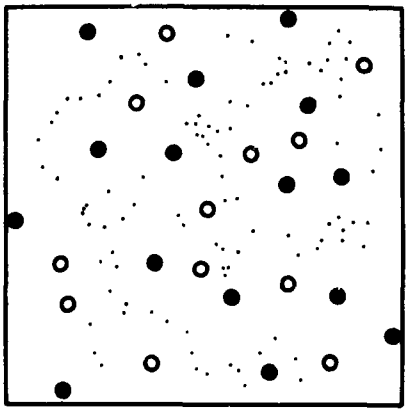


FIGURE 1

Schematic illustration of different defects in a sample. The diffusion length component L_E is due to defects having EBIC contrast (full circles), the component L_M is due to point defect (points) or extended defects giving no EBIC contrast (open circles).

The maximum contrast values c_{max} were near 0.3. The size of the defects showing EBIC contrast was in the range of 1/um so that they could be considered as being small compared to R. The bulk diffusion length L_B was determined from energy-dependent EBIC measurements.

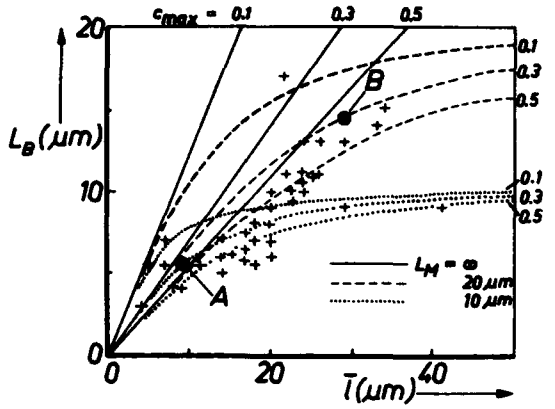


FIGURE 2

Bulk diffusion length L_B in dependence on mean distance l between EBIC contrasts. Measured values (E_o = 30 keV) and calculated curves with EBIC defect contrast c_{max} and diffusion length L_M between the defects as parameters.

The measured data are shown in Fig. 2. More detailed information about two typical samples denoted A and B is given in the table .

It is found that the total defect density obtained by etching (N_{OP} + N_{SF} + ...) is markedly larger than the density of defects giving EBIC contrasts N_E, but N_{SF} and N_E are usually in the same order of magnitude. In sample A $L_E < L_M$ is observed, i.e. the dominant source of bulk recombination are defects showing EBIC contrast, probably SF. Sample B has $L_E \simeq L_M$, again indicating a significant contribution to bulk recombination by defects with EBIC contrast. The diffusion length between the contrast sites, L_M, is estimated to be smaller than the diffusion length in the denuded zone L_{DZ} for both samples. This is thought to be indicative of other but not dominant recombination sources which could exist in the bulk but are missing in the DZ due to IG and which are not detectable by the EBIC method. In the light of our results the OP cannot be therefore considered to play the dominant role in bulk recombination in IG samples.

Dominant OP recombination seems to be rather confined to samples with very high OP densities $N_{OP} = 10^{12}$ ... 10^{13} cm^{-3} /5/ produced by long-duration heat treatments not typical of normal IG procedures.

Nevertheless, oxygen precipitation may be involved in defect recombination in some other respect. It is widely accepted that recombination activity visible as EBIC contrast is due to some interaction of extended crystal defects with impurities. Metallic impurities are considered here, but also oxygen as the most important impurity in Cz silicon. Let us assume that defect recombination activity is simply due to an accumulation of solved metallic impurities at the defect site. For estimating under which conditions such accumulation can produce significant contrast take an active defect volume of 1 μm^3 and an impurity solubility of 10^{16} cm^{-3} which corresponds to 10^4 solved atoms per defect. From (1) one obtains $\sigma_n > c \times 10^{-12}$ cm^2 as an estimate for the minimum capture cross section of the centres necessary for producing the contrast. So, $\sigma_n > 10^{-13}$ cm^2 or even much more would be required for typical contrasts of $c>0.1$, i.e. σ_n hardly expected for simple metallic impurities. Instead, we suggest positive space charges to exist around the defects which support electron capture and increase the apparent cross section of the defect. Maybe oxygen-related precipitates segregating at secondary crystal defects contribute in p-type silicon to these space charges similar to the OP investigated in /5/.

sample	Wright etching		EBIC		calculated using (2) and (3)			
	N_{OP} (cm^{-3})	N_{SF} (cm^{-3})	N_E (cm^{-3})	c_{max}	L_B (/um)	L_{DZ} (/um)	L_M (/um)	L_E (/um)
A	10^{10}	$6x10^9$	$3x10^9$	0.3	5.5	30	11	6.3
B	10^9	10^8	$2x10^8$	0.3	14	3.0	20	20

N_{OP}, N_{SF} ... densities of OP and SF, respectively, determined by Wright etching
$N_E = (1^2R)^{-1} \simeq (1^2.6/\mu m)^{-1}$ ... density of EBIC contrasts at E_o = 30 keV
L_{DZ} ... diffusion length in the DZ determined by the method given in /2/.

Useful information for answering these questions is expected from analytical electron microscopy.

3. INFLUENCE OF DIFFUSION-LENGTH DEPTH PROFILE ON DEVICE PROPERTIES

For substrate-based devices the leakage current j_L of p-n or field-induced junctions at device operation temperature T_o is dominated by its diffusion component /4, 11/.

So, the bulk recombination properties may be important for device operation, e.g. the refresh behaviour of memory devices. The diffusion current density j_d is given by

$$j_d \simeq \frac{q\, n_i^2\, D}{N\, L_S \Phi} \qquad (4)$$

with q being the elementary charge, n_i the intrinsic carrier concentration, D the minority carrier diffusivity, N the dopant concentration, L_S the diffusion length at the surface, and $\Phi < 1$ a correction function depending on the diffusion length profile L(z) /11/. $L_S\Phi$ may be interpreted as effective diffusion length L_S^* at the surface and is easily obtained by EBIC measurements /11, 2/. Typically, L_S^* is in the 10 μm range for IG wafers /11/.Using L_S^* (T_R) measured at room temperature T_R, it is possible to estimate the leakage current $j_L(T_o)$ at T_o /11/.

The relationship between L(z) and latch-up immunity is discussed in /3/.

REFERENCES

/1/ H. Richter, Proc. 1st. Internat. Autumn School GADEST 1985, Academy of Science of GDR, Garzau, Ed. H. Richter (p. 1)

/2/ C. Donolato and M. Kittler, submitted to J. Appl. Phys.

/3/ M. Kittler, W. Seifert ,and C. Donolato, Proc.GADEST 1987,in print

/4/ D.K. Schroder, Solid State Electronics 27 (1984) 247

/5/ J. M. Hwang and D.K.Schroder, J. Appl. Phys. 59 (1986) 2476

/6/ C. Donolato, Optik 52 (1978/79) 19

/7/ L. Pasemann, H. Blumtritt ,and R. Gleichmann, Phys. Stat. Solidi (A) 70 (1982) 197

/8/ M. Kittler and W. Seifert, Phys. Stat. Solidi (A) 66 (1981) 573

/9/ N. Inoue, C.W. Wilmsen, and K.A. Jones, Solar Cells 3 (1981) 35

/10/ T.E. Everhart and P.H. Hoff, J. Appl. Phys. 42 (1971) 5837

/11/ M. Kittler and W. Seifert, Phys. Stat. Solidi (A) 99 (1987) 559

CVD AND I.C. METALLIZATION

Jan-Otto CARLSSON

Institute of Chemistry, Solid State Chemistry Group,
University of Uppsala, P.O. Box 531, S-751 21 Uppsala, Sweden

1. INTRODUCTION

The feature size of integrated circuits are continuosly scaled downwards. This opens requirements of new materials as well as new process technologies. For metallization in VLSI, materials having low resistivity, low contact resistance and high electromigration resistance are of interest. Additional materials requirements are that they are easily etched and can act as diffusion as well as etching barriers.

With the use of shallower juctions unwanted dopant diffusion and redistribution become more critical. Hence there is a need of low-temperature deposition technologies. For the interconnect metallization, an excellent step coverage, reducing electromigration effects, is required. Finally, with the decreased alignement tolerance of the successive mask levels, self-aligning deposition technologies are attractive.

For VLSI applications the refractory metals (in particular W, Mo and Ti) as well as their disilicides are of interest. The CVD technique is known to yield excellent step coverage and conformality, good uniformity, high purity of the films deposited, and the possibility to selective deposition. The attractive metallization properties of tungsten and the rapid progress in selective tungsten CVD (a self-aligning technique) focus this paper to CVD of tungsten. Selective deposition of tungsten is regarded to be a highly useful technique in the very large scale and ultralarge scale IC fabrication (1,2). Tungsten has been proposed to be used in for instance Schottky diode barriers (3), current shunts (1,4), diffusion barriers (4), contact and via fill for planarity (5).

This paper consists of two parts:

i) A general introduction to CVD with emphasis to trends in the design of CVD reactors and to optimization of the throwing power of a CVD process.

ii) Tungsten CVD on silicon and disilicides, respectively, including process modeling, mechanisms, interfacial reactions, and finally the relationship between electrical characteristics and deposition conditions.

Space limitations preclude presenting details. For details the reader is referred to the references given.

2. CHEMICAL VAPOR DEPOSITION

2.1. CVD reactors

In CVD a solid material is obtained by a chemical reaction between gaseous reactants. Reactions in the vapor as well as on a substrate surface may be combined to yield the deposited material. Two main reactor types can be distinguished:

In the hot wall reactor the reactor tube is surrounded by a tube furnace. This means that the substrates and the wall of the reactor have the same temperature. In this reactor type, many wafers can be coated in the same run. But because of the high temperature of the reactor walls, the deposition occurs not only on the substrates but also on the reactor walls. This

may result in a depletion of the vapor with respect to the reactants - changes in the deposition rates - when the reactants are transported through the reactor. There is also a risk of introducing contaminants in the system from chemical reactions between the reactor wall and the vapor. Other contaminants may be particles, which have loosed from the reactor wall.

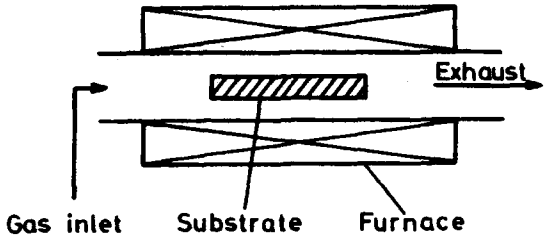


FIGURE 1

Hot wall reactor

In the cold wall reactor the substrates have a higher temperature than the reactor wall, which means that the deposition occurs only on the substrates. The risk of contamination, originating from an interaction between the reactor wall and the vapor, is considerably reduced in this reactor type. However due to the steep temperature gradient around the substrates severe natural convection can arise. Finally the depletion of the vapor with respect to the reactants is normally less in a cold wall reactor than in a hot wall reactor.

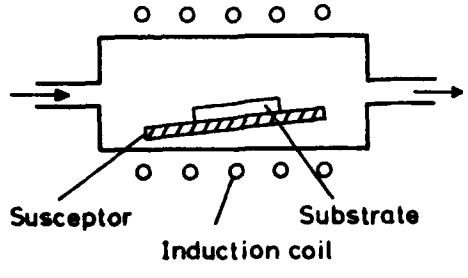


FIGURE 2

Cold wall reactor

In the processing of devices with geometries below 2 μm, particle contamination may be a factor, controlling the device yields. In the hot wall reactor, deposition also occurs on the inside of the reactor walls. With increasing thickness of the deposit, the contamination risk from particles, loosing from the reactor walls, are apparent. Hence there is a trend (at present) to prefer cold wall reactors in processes not controlled by homogeneous reactions in the vapor. For further contamination control as well as reduced process time, loadlock capabilities may be used. Finally, CVD systems dedicated for only one process, can be constructed in materials compatible with an etching gas for cleaning of the chamber.

2.2. Rate-limiting steps

In CVD the substrates are immersed in a gas stream, resulting in the development of so called boundary layers. The boundary layers are defined as the region near the substrate surface where the gas stream velocity (velocity boundary layer), the concentration of the different vapor species (concentration boundary layer) and the temperature (thermal bounday layer) are not equal to those in the main gas stream.

In a CVD process various sequential reaction steps occur. Each of these steps may be rate-limiting in the absence of thermodynamic limitations. Plausible rate-limiting steps in a CVD-process are listed below (6-9):

a) Transport of the gaseous reactants to the boundary layer surrounding the substrate (free and forced convection).
b) Transport of the gaseous reactants across the boundary layer to the surface of the substrate (diffusion and convection flows).
c) Homogeneous reactions in the vapor during the transport to the substrate surface.
d) Adsorption of the reactants on the surface of the substrate.
e) Chemical reactions (surface reactions) between adsorbed reactants or between adsorbed reactants and reactants in the vapor.

f) Nucleation.
g) Desorption of some of the reaction products from the surface of the substrate.
h) Transport of the reaction products across the boundary layer to the bulk gas mixture.
i) Transport of the reaction products away from the boundary layer.

In each of these steps several processes may proceed simultaneously.

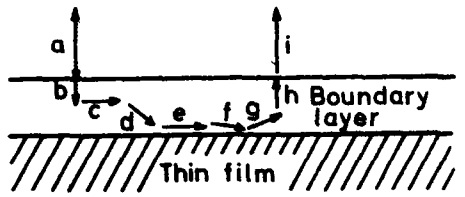


FIGURE 3

The sequential steps in CVD.

Even though several rate-limiting steps can be identified in a CVD process, only four main categories of control are normally discussed:

i) *Thermodynamic control:* Thermodynamic control means that the deposition rate is equal to the mass input rate into the reactor (corrected for the yield of the process). This occurs at extreme deposition conditions (low gas flow rates, high temperatures, ...). The temperature dependence of the deposition rate is obtained from thermodynamic calculations.
ii) *Surface kinetics control:* If the deposition rate is lower than the mass input rate into the reactor as well as the mass transport rate in the vapor to or from the subatrate surface, a surface kinetics control or nucleation control exist.
iii) *Mass transport control:* A process may also be controlled by the mass transport in the vapor to or from the substrate surface. This occurs frequently at high pressures and high temperatures and when very instable compounds are used as reactants.
iv) *Nucleation control:* At low supersaturation the deposition rate may be controlled by the nucleation.

The four types of control as defined above are those normally discussed. However recent investigations indicate (see for instance refs. /10,11/) that also homogeneous reactions in the vapor play an important role in CVD and that they may control the deposition rate. So in addition to the four controls given above homogeneous reaction control should be introduced.

2.3. Experimental conditions for maximum in step coverage

With increasing aspect ratios in the geometries fabricated, experimental conditions for maximum step coverage, conformality and uniformity have to be determined. This maximum occurs at surface kinetics control, which may be achieved in different ways:

i) By increasing the gas flow velocity. The deposition rate at surface kinetics control is given by the limit value.

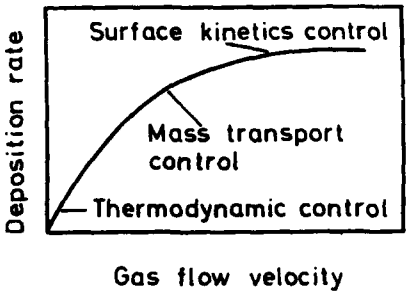


FIGURE 4

Influence of gas flow velocity on the control of a CVD process

ii) By decreasing the temperature and/or decreasing the total pressure. For an endothermic CVD process the surface kinetically controlled region is characterized by a high activation energy, while the mass transport controlled region has a much lower activation energy. As can be seen from the figure below, the surface

kinetically controlled region is expanded at lower total pressures, which is due to a higher diffusivity of the molecules in the vapor at lower total pressures.

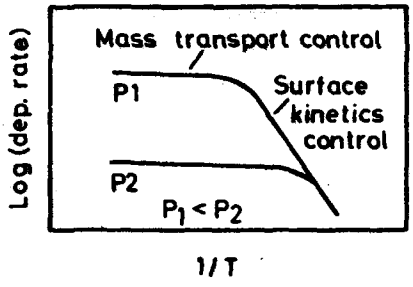


FIGURE 5

Influence of temperature on the control of a CVD process

iii) By increasing the thermochemical stability of the reactants used in the process. By using $SiCl_4$ instead of SiH_4 in silicon CVD, the surface kinetics controls the deposition rate at higher temperatures.

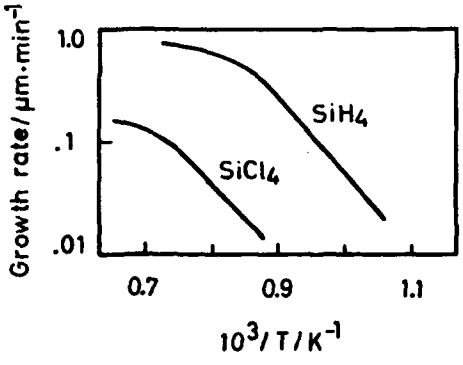


FIGURE 6

Influence of the thermochemical stability of the source molecules on the control of a CVD process

In conclusion for a particular gas mixture maximum in step coverage can be expected at low total pressures, low temperatures and high gas flow velocities.

3. TUNGSTEN CVD ON SILICON AND DISILICIDES

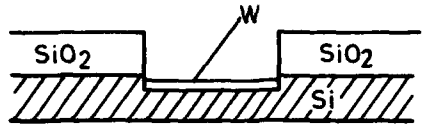
In CVD of tungsten, WF_6 and H_2 are used as reactants. In the selective tungsten deposition process on silicon (see fig. 7), two main reaction steps can then be distinguished:

i) In the first step elemental silicon will act as the predominating reducing agent. This results in tungsten deposition on those substrate regions where elemental silicon is exposed to the vapor.

ii) In the second step another reducing agent, hydrogen, has to take over, since the tungsten film obtained in the first step, separates the elemental silicon from the vapor.

STEP I

$$3Si(s) + 2WF_6(g) \longrightarrow 2W(s) + 3SiF_4(g)$$



STEP II

$$3H_2(g) + WF_6(g) \longrightarrow W(s) + 6HF(g)$$

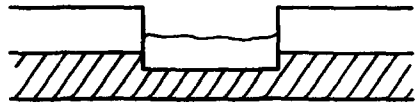


FIGURE 7

The different process steps in selective tungsten CVD

Before discussing details of the two process steps, a general discussion of selective tungsten deposition, based on thermodynamics, will be given.

3.1. Thermodynamic analysis of selective tungsten deposition

Thermodynamics can be used as a guide for prediction of trends in selectivity when the deposition conditions are changed. It can also be used to identify plausible (and undesired)

side reactions as well as gaseous selectivity modifiers (12). The trends in selectivity in the two process steps should be analyzed from different calculated thermochemical parameters. For step one, calculated driving force values for the reactions between elemental silicon and the vapor and driving force values for reactions between silicon dioxide and the vapor, respectively should be used. In the following some results from thermochemical investigations of selective tungsten deposition are summarized. Free energy minimization technique was used in the calculations and about 100 different substances were included in the calculations. For details, the reader is referred to ref. (13).

Table 1. Calculated driving force (kJ/mol reaction gas mixture) for the reaction between silicon and the vapor, $\Delta G(Si)$ and the reaction between silicon dioxide and the vapor. $\Delta G(SiO_2)$, respectively.

H_2/WF_6	T/°C	P/torr	$\Delta G(Si)$	$\Delta G(SiO_2)$
39	300	0.1	-19.5	-7.4
39	500	0.1	-20.1	-9.4
39	300	1.0	-19.3	-7.2
39	500	1.0	-19.9	-8.9

Examples of calculated driving force values are given in table 1. As can be seen a much higher driving force is obtained for the Si/vapor reaction than for the SiO_2/vapor reaction. This indicates a preference for deposition of tungsten to silicon substrate areas. Hence a self-aligned process has been obtained. The vapor species formed at the Si/vapor reaction is shown in figure 8. With the formation of silicon fluorides silicon has been etched. For stoichiometric reasons, a tungsten layer of 100 Å will consume about 200 Å silicon. From table 1 it can also be deduced that a low temperature, a high total pressure and a high H_2/WF_6 molar ratio should favor selective growth in the first step. Finally, in the SiO_2/vapor reaction tungsten oxides and different gaseous substances may be formed.

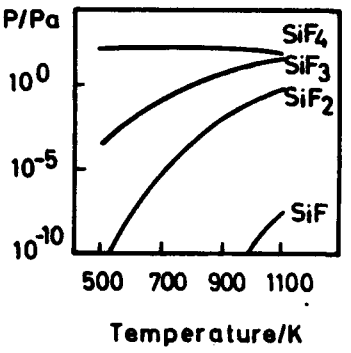


FIGURE 8

Partial pressures of the vapor species formed in the reaction between silicon and tungsten hexafluoride. Total pressure = 0.1 kPa (13).

The thermodynamic analysis of step one indicated the risk of tungsten oxide formation. Thus the selectivity in step two should be analyzed from the assumption of formation of tugnsten oxides, i.e., the results from such an analysis indicates the possibility to maintain selectivity even if tungsten oxides have nucleated on the silicon dioxide. Calculated chemical potentials for homogeneous equilibrium in the vapor (no solid substance present), equilibrium between the vapor and the silicon and equilibrium between the silicon dioxide and the vapor, respectively for various temperatures are shown in figure 9. The driving force for tungsten deposition in the "silicon window" or the driving force for tungsten oxide formation is given by the separation of the "homogeneous" curve and the curves for deposition in the "silicon window" and deposition on the tungsten oxides. As can be seen from the diagram, selectivity to the silicon window is favored by a higher temperature. Similar analysis for other total pressures and vapor compositions indicate that selective deposition to the silicon window is favored by a low total pressure and a high H_2/WF_6 molar ratio.

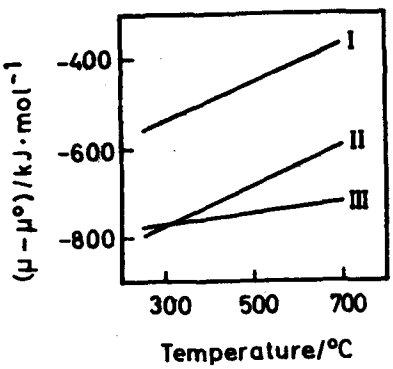


FIGURE 9

Chemical potential of tungsten in the vapor, μ, minus a reference potential, μ^0, at different temperatures. H_2/WF_6=40. total pressure = 0.1 torr (12).

Defining a selectivity number as the driving force for the reaction between the material in the silicon window devided by the driving force for the reaction between the "silicon dioxide" region and the vapor, a useful parameter for discussion of selectivity has been obtained. A selectivity number larger than one indicates preference for deposition to the silicon window. As an example, the influence of temperature on the selectivity number is shown in figure 10.

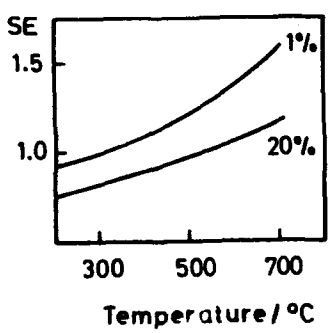


FIGURE 10

Calculated selectivity number as a function of temperature for two different H_2-WF_6 gas mixtures (1% WF_6 and 20% WF_6, respectively). Total pressure: 0.1 torr (12).

For optimization of selectivity in general in deposition processes, gaseous selectivity modifiers can be added to the reaction gas mixture. These modifiers influence in particular the homogeneous reactions in the vapor. For improvement of the selectivity in tungsten CVD, a stronger reducing agent than hydrogen should be added to the H_2/WF_6 gas mixture. An example of such an agent is SiH_4. However, addition of SiH_4 to the actual gas mixture may cause silicide deposition. From the calculated CVD phase diagram (see fig. 11) it can be seen, that relatively large concentrations of SiH_4 is required for silicide growth. Thus SiH_4 may be a useful selectivity modifier.

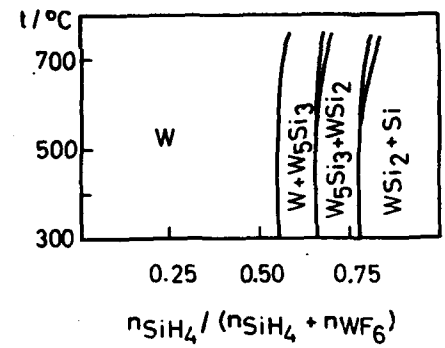


FIGURE 11

Calculated CVD phase diagram for the system W-Si H_2/WF_6 = 39, total pressure = 0.1 torr. (12).

Calculated selectivity number as a function of SiH_4 concentration in the vapor is shown in figure 12. A drastic change in the selectivity number, with expected improvement in selectivity, is obtained at a particular SiH_4 concentration. The selectivity is improved by a factor of about 10 (12). This change in the selectivity is mainly due to the opening of a new reaction channel in the vapor at a particular SiH_4 concentration.

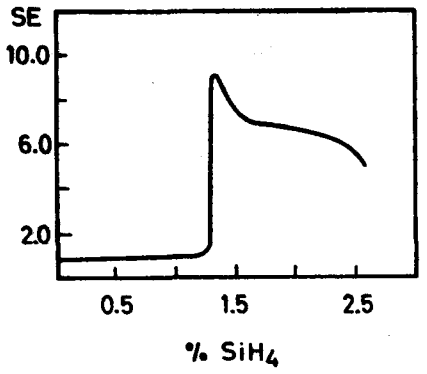


FIGURE 12

Calculated selectivity number as a function of SiH_4 concentration. Temperature = 300 °C, total pressure = 0.1 torr. H_2/WF_6 = 39 (12).

3.2. Tungsten CVD; Process aspects and materials properties

Selective deposition processes, based on vapor/solid interfacial reactions, are highly sensitive to the surface conditions of the whole substrate surface. In selective tungsten deposition, silicon as well as the adjacent silicon dioxide should be properly prepared. An oxide with a low defect and impurity content in combination with silicon, free from contamination and native oxide, creates good selectivity.

3.2.1. Silicon reduction of WF_6

The initial stage in tungsten CVD was controlled by the silicon reduction. A self-limiting growth process, due to separation of the silicon from the vapor, was obtained. The thickness of the tungsten layer after the self-limiting reaction is affected by the cleaning procedure prior to the deposition and the residual native oxide thickness, total pressure, deposition temperature and the doping of the silicon (14,15). For standard deposition conditions (300°C, H_2/WF_6=40, total pressure 0.5 torr) and a native oxide thickness less than 10 Å, the thickness of the tungsten layer is about 100 Å. Busta and Tang (14) have reported an increase in the tungsten layer thickness with increasing native oxide thickness up to about 20 Å. For a thicker native oxide, the thickness of the tungsten layer decreases drastically (see fig. 13).

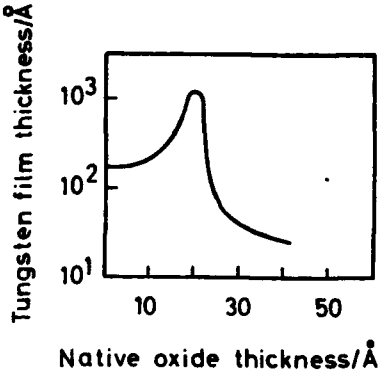


FIGURE 13

Tungsten film thickness as a function of the native oxide film thickness. Data from ref. (14).

The initial reaction consumed silicon. This results also in a lateral growth of tungsten under SiO_2. It has been shown, however, that this lateral growth can be considerabel with an encroachment of tungsten under SiO_2, yielding a source of junction leakage currents. This effect is probably due to the variation in the thickness of the native oxide across the silicon window. Near the SiO_2 sidewalls a thicker oxide can be expected. According to the discussion in the paragraph above, an increased consumption of silicon, yielding tungsten encroachment, can be expected at thicker native oxides. By a low WF_6 partial pressure, this encroachment can be eliminated (16).

Formation of tunnels (~100 Å in diameter) has been observed by several investigators (see for instance ref. 17). They may be initiated at surface sites, where tungsten aggregates, particles, are generated. These particles may act as a getter for atomic fluorine, supporting local etching of silicon. The tunnels can be avoided by a proper choice of deposition conditions (18).

3.3.2. Hydrogen reduction of WF_6

The second step in the selective tungsten deposition is based on the fact that growth of a material can be maintained at a low supersaturation, while heterogeneous nucleation requires a higher supersaturation. This means that for low supersaturation, growth on already deposited material will occur and no nuclei will be formed outside the region of the deposited material (for instance on the oxide). In tungsten CVD, a low supersaturation means a few per cent WF_6. By using a low supersaturation tungsten layers with thicknesses up to 1 μm have been grown selectively (19). Defects in the oxide as well as contaminants will degrade the selectivity.

3.2.3. Kinetics and mechanisms

The initial reaction is very fast and within a few seconds the self-limiting reaction is completed. The deposition rate is much lower in step two than in step one. Typical growth rates are shown in figure 14. As can be seen, an increase in the temperature from 300 to 350°C increases the deposition rate in step two roughly with a factor of three.

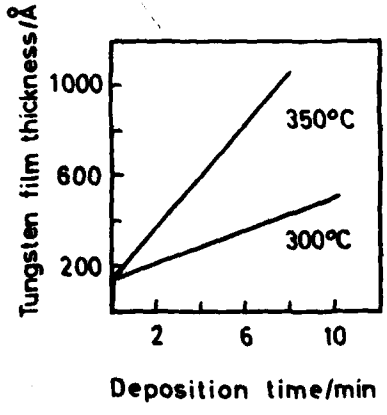


FIGURE 14

Tungsten film thickness as a function of the deposition time. H_2/WF_6=15, total pressure 0.5 torr, total flow = 17000 cm³/min. Data from ref. (20).

In step one, the deposition mechanism is a displacement reaction. For step two it has been observed that the deposition rate is independent of the WF_6 partial pressure but dependent on the H_2 partial pressure (square root dependence). The influence of the H_2 partial pressure on the deposition rate indicates that the dissociation of the H_2 molecules may be rate-limiting (20). The HF formed during the process decreases the deposition rate. This can be explained in terms of a Langmuir-Hinschelwood mechanism (21). A large HF concentration in the vapor, which is obtained at large tungsten areas exposed to the vapor, slows down the deposition rate considerably and degrades the selectivity (21).

3.2.4. Materials characterization

Two polymorphs of tungsten exist: α-tungsten and β-tungsten. β-tungsten is stabilized by small amounts of oxygen and fluorine. β-tungsten has a resistivity about an order of magnitude higher than α-tungsten. α- as well as β-tungsten has been observed at tungsten CVD on silicon. β-tungsten is formed at thicker native oxide layers (~20Å) (14). No silicide formation has been observed.

Films of α-tungsten are free from texture (22). The grain size seems to be independent of the deposition temperature but dependent on film thickness (see fig. 15). The oxygen concentration is influenced by the deposition temperature as well as the film thickness. At a film thickness of 200 nm, the oxygen concentration is about 0.02 and 0.07 at.% at 300 and 400°C, respectively. The oxygen is segregated to the grain boundaries. In contradiction to the oxygen concentration, the fluorine concentration in the films decreases with increasing temperature (at 400°C ~0.03 at.% F) (23).

The tungsten films exhibit tensile stresses. These stresses are influenced by the film thickness. For thin films, stresses up to $1.3 \cdot 10^{10}$ dyn/cm² have been observed. A minimum in stress is obtained at a film thickness of about 500 Å.

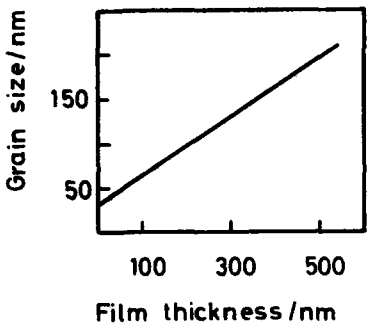


FIGURE 15

Grain size as a function of tungsten film thickness. Data from ref. (23).

For thicker films the stress increases gradually and is about $8 \cdot 10^9$ dyn/cm^2 for a thickness of 2000 A (22).

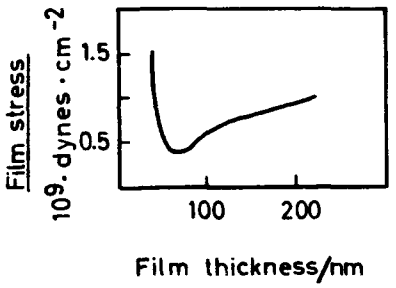


FIGURE 16

Film stress as a function of tungsten film thickness. Data from ref. (22).

3.2.5. Electrical characterization

The resistivity of CVD tungsten films is shown in figure 17. The resistivity increases drastically for film thicknesses less than 500 Å. In thicker films resistivity values close the bulk value (5.3 $\mu\Omega \cdot$cm) is obtained. The thickness dependence of the resistivity can be explained in terms of scattering by grain boundaries, and impurities (23).

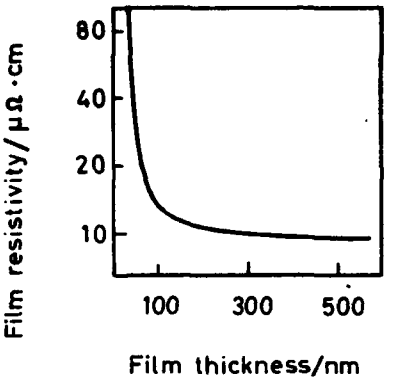


FIGURE 17

Film resistivity vs. film thickness. Data from ref. (22).

3.2.6. Tungsten CVD on disilicides

The tunnel formation in the silicon as well as the tungsten encroachment underneath the silicon dioxide may cause reliability problems. Deposition conditions for elimination/reduction of these problems have been discussed above. An alternative metallization scheme may use an interlayer in between tungsten and silicon. The tungsten is then used as a diffusion barrier and/or a filling material.

Use of interlayers may be associated to a new set of problems. In this section, a general introduction to three major problems at CVD of tungsten from H_2/WF_6 on interlayer materials will be given. Since disilicides are of particular interest, they will be used as examples.

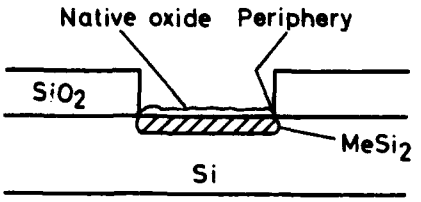


FIGURE 18

Problem areas when using interlayer technique

The three major problems are the following (see fig. 18):

- The chemical composition of the oxide, covering the silicide,
- the difference in chemical reactivity between the elements composing the silicide,
- the reaction between silicon and the vapor at the periphery of the contact hole.

$TiSi_2$ is used to illustrate the first two problems, while use of $CoSi_2$ illustrates the third problem.

The surface oxide on $TiSi_2$ contains both Ti and Si. No chemical etch for removal of this oxide exist. With an HF dip for instance, Si is preferentially etched and a more Ti rich surface oxide is obtained. Thus at tungsten CVD, the $TiSi_2$ substrate exposes an oxide to the vapor. This oxide reacts with the H_2/WF_6 vapor. At a low temperature, solid TiF_3 is formed. The remaining oxide becomes then more Si rich. For a higher temperature, Si in the SiO_2 is etched away and a more Ti rich oxide together with tungsten oxides can be expected to be formed (24).

Since the native oxide on $TiSi_2$ varies in thickness and has defects, it will be penetrated here and there with a subsequent reaction between $TiSi_2$ and the vapor. At this reaction different reaction products may be formed. From the calculated CVD phase diagrams (see fig. 19) it can be seen, that solid TiF_3 may be formed at a low temperature (500 K). Silicon in $TiSi_2$ is also etched preferentially at higher H_2/WF_6 molar ratios and lower total pressures, resulting in the formation of TiSi.

From the brief discussion of the interfacial reactions between $TiSi_2$ and the H_2/WF_6 vapor, the following can be concluded: Due to the non-uniform penetration of the surface oxide on $TiSi_2$, a rough interface between tungsten and $TiSi_2$ can be expected. This interface contains oxygen (from unreacted surface oxide) and solid TiF_3. This has also been verified experimentally (24,25). The formation of TiF_3 results in poor adhesion and high contact resistance.

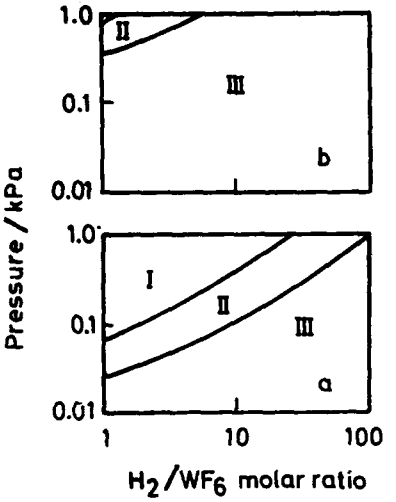


FIGURE 19

Solid phases in equilibrium with $TiSi_2$ (excess of $TiSi_2$ assumed in the calculations): I = TiF_3 + W + Si, II = TiF_3 + W, III = TiSi + W a) 500K, b) 550K (24).

Tungsten can be deposited selectively on $CoSi_2$ in a surface catalyzed reaction (26). In the periphery of the contact, however, a reaction between silicon and the H_2/WF_6 vapor occurs, i.e. $CoSi_2$ doesn't separate the silicon from the vapor effectively. This reaction results in the formation of voids at the SiO_2/Si interface as well as encroachment of tungsten underneath the SiO_2 (26).

The formation of TiF_3 in the $TiSi_2/W$ interface caused poor adhesion and high contact resistance. Voids and tungsten encroachment in the Si/SiO_2 interface are obtained in self-aligned $CoSi_2$ contact structures. A low contact resistance have been reported for $W/MoSi_2$ and $W/TaSi_2$ interfaces (5, 27).

4. CONCLUDING REMARKS

CVD is an attractive technique for VLSI processing, since high-purity films of good uniformity and step coverage can be grown. The

selective deposition possibilities are of highest interest when the feature size of integrated circuits is reduced.

Selective deposition processes are sensitive to the cleaning as well as the defects of the substrate material. Hence proper substrate pretreatment procedures have to be developed.

Selective CVD processes are based on differences in interfacial chemistry between different substrate areas. These differences may be favorably enhanced by a proper change in the composition of the vapor, the deposition temperature or by the total pressure.

Tungsten CVD has many advantages in VLSI processing and will certainly be applied in some process steps. The present problems of this process (tunnel formation, tungsten encroachment in the Si/SiO_2 interface) limits its use today to multilevel via fill and interconnect/diffusion barrier plug.

For elimination/reduction of undesired interfacial reactions between the H_2/WF_6 vapor and the substrate, yielding high contact resistance, poor adhesion, voids, ... the interlayer technique is promising. The different process steps (preparation of interlayer, substrate pretreatment, tungsten CVD) should be combined and designed for optimum in adhesion, electrical characteristics and reliability.

REFERENCES

[1] Gargini, P.A. and Beinglass, I., IEDM Tech. Dig., 54 (1981).

[2] Broadbent, E.K. and Ramiller, C.L., J. Electrochem. Soc., 131 (1984) 1427.

[3] Crowell, C., Savace, J. and Sze, S., Trans. Metall. Soc. AIME, 233 (1965) 478.

[4] Miller, N.E. and Beinglass, I., Solid State Technol., 25 (1982) 85.

[5] Moriya, T., Shima, S., Hazuki, Y., Chiba, M. and Kashiwagi, M., IEDM Tech. Dig., CH 1973-7/83 (1983) 550.

[6] Faktor, M.M. and Garret, I., Growth of Crystals from the Vapor, Chapman and Hall, London, (1974).

[7] Battat, D., Faktor, M.M. and Garret, I., in W.R. Wilcox (ed.), Preparation and Properties of Solid State Materials, Vol. 2. Marcel Dekker, New York (1976).

[8] Shaw, D.W., in C.H.L. Goodman (ed.), Crystal Growth - Theory and Technique, Vol. 1, Plenum Press, New York (1974).

[9] Jones, M.E. and Shaw, D.W., in M.B. Hannay (ed.), Treatise in Solid State Chemistry, Vol. 5, Plenum Press, New York (1975).

[10] Coltrin, M.E., Kee, R.J. and Miller, J.A., Proc. of the 9th Int. Conf. on CVD, eds.: Mc D. Robinsson, van den Brekel, C.H.J., Culten, G.W., Blocher Jr., J.M., and Rai-Choudhury, P., The Electrochem. Soc., Pennington, N.J. (1984) 31.

[11] Lindström, J.N. and Stjernberg, K.G., Proc. of the 5th European Conf. on CVD, eds.: Carlsson, J.O. and Lindström, J.N. (1985) 169.

[12] Carlsson, J.O. and Hårsta, A. To be published.

[13] Carlsson, J.O. and Boman, M., J. Vac. Sci. and Technol., A3(6) (1985) 2298.

[14] Busta, H.H. and Tang, C.H., J. Electrochem. Soc., 133(6) (1986) 1195.

[15] Taso, K.Y. and Busta, H.H., J. Electrochem. Soc., 131(11) (1984) 2702.

[16] Moriya, T., Yamada, K., Shibita, T., Fizuka, H. and Kashiwai, M., Proc. 1983 Symp. VLSI Technol., Maui, HI, (1983) 96.

[17] Stacy, W.T., Broadbent, E.K., Norcott, M.H. J. Electrochem. Soc., 132(2) (1985) 444.

[18] Pauleau, Y., Lami, Ph., Tissier, A., Pantel, R. and Oberlin, J.C., Thin Solid Films, 143 (1986) 209.

[19] Blewer, R.S., Tracy, M.E. and Wells, V.A., Tungsten and Other Refractory Metals for VLSI Applications, ed. Blewer, R.S., Materials Research Society, Pittsburgh, PA, (1986) 21.

[20] Broadbent, E.K. and Ramiller, C.L., J. Electrochem. Soc., 131 (6) (1984) 1427.

[21] Pauleau, Y. and Lami, Ph., J. Electrochem. Soc., 132 (11) (1985) 2779.

[22] Green, M.L. and Levy, R.A., J. Electrochem. Soc., 132 (5) (1985) 1243.

[23] Learn, A.J. and Foster, D.W., J. Appl. Phys., 58(5) (1985) 2001.

[24] Eriksson, Th., Carlsson, J.O., Niemi, E., Östling, M. and Petersson, C.S. To be published.

[25] Broadbent, E.K., Morgan, A.E., DeBlasi, J.M., van der Putte, P., Coulman, B., Burrow, B.J., Sadana, D.K. and Reader, A., J. Electrochem. Soc., 133 (8) (1986) 1715.

[26] van der Putte, P., Sadana, D.K., Broadbent, E.K. and Morgan, A.E., Appl. Phys. Lett., 49(25) (1986) 1723.

[27] Draper, B.L., Hill, T.A. and Bell, H.B., Proc. 2nd Int. IEEE VLSI Multilevel Interconnect. Conf., IEEE Cat. No. 85 CH 2197-2, (1985) 90.

Session A2.2

Ultrafast Bipolar II

Chairman: L. Treitinger

Tuesday, September 15, 1987

Self-Aligned Technology for Sub-100nm Deep Base Junction Transistors

Masahiko Nakamae

1st LSI Division, NEC Corporation
1120, Shimokuzawa, Sagamihara, Kanagawa, Japan

The problems in scaling down of modern advanced polysilicon self-aligned transistors are briefly discussed. Then, a novel self-aligned technology is proposed to solve the problems. The newly developed BSA (BSG Self-Aligned) technology is featured by the use of CVD-BSG film as a sidewall spacer as well as a diffusion source to form both intrinsic base and p^+-connecting regions, simultaneousely. The fabricated transistor having 40nm deep emitter-base junction and sub-100nm collector-base junction shows 70 of h_{FE} and 7V of BV_{CEO}, respectively.

1. Introduction

It is well known that the recent progress in high speed performance of Si bipolar transistors has been mainly achieved by the use of modern advanced polysilicon self-aligned technologies [1],[2]. Using these technologies, the transistor dimension has been scaled down to an ultimately small size, and the drastic reduction of the parastic capacitance and resistance has been performed.

For further progress, it is indispensable to form very shallow base region to achieve higher f_T value. But, conventional ion implantation method faces two problems in reducing the base junction depth. The first has to do with the fact that it is quite difficult to obtain sub-200nm deep base junction because of the secondary channeling effect of boron ions even at very low acceleration energy. The second problem has to do with the fact that it becomes impossible to anneal out the lattice defects formed by ion implantation at a higher dose range.

In addition, it is not possible to solve the inherent problem of the advanced self-aligned transistor like SST[1], relating to the p^+-connecting region between emitter and extrinsic base regions.

2. BSA process

The main process steps in fabricating the BSA transistor are schematically shown in Fig. 1.

(a) The emitter window was opened in the SiO_2/p^+-polysilicon stacked layer deposited on the oxide isolated epitaxial layer. After the drive-in process to form the extrinsic base region, BSG film was deposited on the entire surface.

(b) Rapid thermal annealing (RTA) was performed to form the sub-100nm deep intrinsic base region.

(c) The BSG film was directionally etched using reactive ion etching technique.

(d) The extremely shallow emitter was formed by arsenic diffusion from heavily ion-implanted polysilicon using RTA again. In this step, intrinsic base and the connecting regions received RTA treatment once more. But, the boron profile was given approximately by Gaussian distribution in the intrinsic base region, and by complementary error function, in the connecting region respectively.

As the result, the optimization of the boron profiles in both regions was successfuly performed by controlling the boron doping concentration in the BSG film and RTA conditions.

3. Result and Discussion

Fig. 2 shows the boron profiles formed by the BSA technology, and low energy ion-implantation method for comparison. In the ion-implantation case, the tailing charachteristics was clealy observed and the junction depth at $10^{16}cm^{-3}$ concentration was about 220nm. But, in the BSA case, extremely shallow profile was achieved, 60nm to 120nm, by changing RTA time, 10sec to 60sec at 1000°C.

Fig. 3 shows the h_{FE} and BV_{CEO} of the fabricated BSA transistor as a function of boron concentration in the BSG film. RTA was carried out at a condition of 1000°C-10sec for both the base and emitter regions. At the boron concentration of 4mol% in the BSG film, BSA transistor showed 70 of h_{FE} and 7V of BV_{CEO}, respectively. No interfacial films between the emitter polysilicon and epitaxial layer was used in this experiment. Although the base width of the transistor was about 40nm, good DC charachteristics were obtained due to the high doping concentration of boron. The lattice defects in the base region have been usually observed in such a high doping condition as $10^{19}cm^{-3}$ in the case of ion-implantation method. But, in the case of this BSA technology, no lattice defect was observed, and the "piping" failure was not observed.

Recently, several new approaches have been reported to form such a shallow base [3], it is not possible to solve the design problems relating to the p^+-connecting region. One solution for the problems has been shown by using LDD MOSFET-like process technology [4]. However, it is not applicable to the SST type transistors. In addition, the defect generating process condition cannot be avoided in ion-implantation method. Therefore, the potential of the LDD MOSSFET-like process seems to be small for the future scaled bipolar devices.

4. Summary

A novel self-aligned technology, BSA technology, has been developed to solve the severe problems of future scaled bipolar transistors. Using this technology, the advanced self-aligned transistor having sub-100nm base junction has been successfully fabricated.

References

[1] T. Sakai et al, "Gigabit logic bipolar technology: advanced super self-aligned process technology," Electron. Lett., vol. 19, pp. 283-284,1983.

[2] K. Washio et al, "A 48ps ECL in a Self-Aligned Bipolar Technology," in ISSCC Dig. Tech. papers, pp.58-59, 1987.

[3] H. Park et al, "High-Speed Self-Aligned Polysilicon Emitter/Base Bipolar devices Using Boronand Arsenic Diffusion Through Polysilicon," in Extended Abstracts 18th Int. Conf. on Solid State Devices and Materials, pp. 729-731, 1986.

[4] D. D. Tang et al, "Design Considerations of High-Performance Narrow Emitter Bipolar Transistors," IEEE Trans. Electron Device Letters, vol. EDL-8, pp. 174-175, 1987.

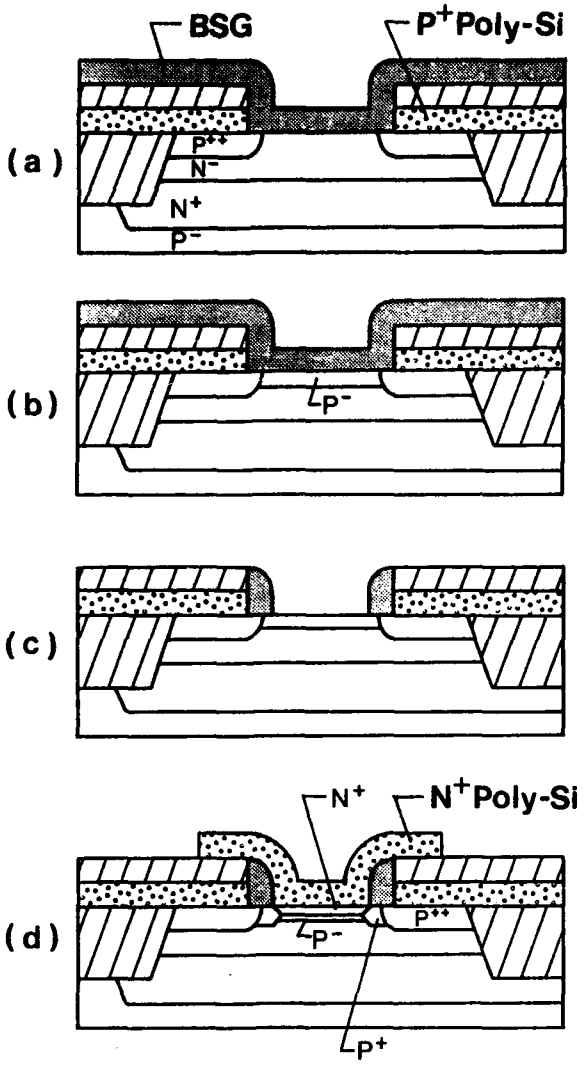


Fig. 1. Main fabrication steps of BSA transistor.

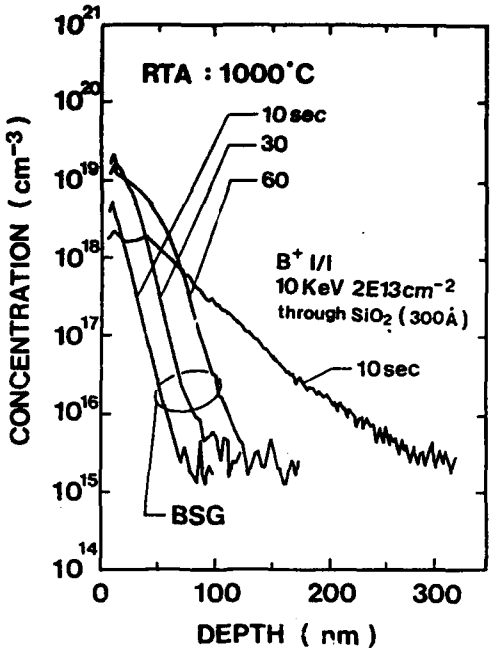


Fig. 2. SIMS depth profiles of intrinsic base formed by BSA technology and low energy ion implantation method.

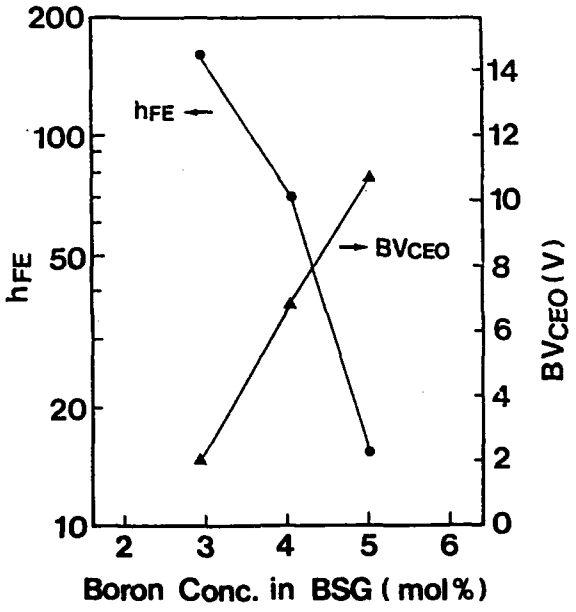


Fig. 3 DC characteristics, h_{FE} and BV_{CEO}, of BSA transistor as a function of boron concentration in BSG film.

VERTICAL SCALING CONSIDERATIONS FOR POLYSILICON-EMITTER BIPOLAR TRANSISTORS

H. Schaber, J. Bieger, B. Benna and T. Meister

Siemens AG, Central Research and Development, Microelectronics,
Otto-Hahn-Ring 6, D-8000 München 83, FRG

1. INTRODUCTION

In contrast to MOS devices, the intrinsic speed of high performance bipolar transistors is limited by vertical dimensions, mainly of the base and emitter regions. Therefore, despite the importance of lateral scaling and self-alignment techniques for reduction of parasitic elements, vertical scaling of dopant profiles remains a key issue for improvement of bipolar device performance. Based on experimental results and on device simulation, this paper discusses present status and possible limitations for vertical scaling of npn transistors with polysilicon emitters.

2. POLYSILICON EMITTERS

Polycrystalline silicon is now widely used in bipolar technology as means of making contact to the emitter and base of the transistor. To minimize oxide contamination for the critical emitter contact usually a HF-dip etch is performed prior to inserting the sample into the polysilicon deposition system.

At low and intermediate current levels the polysilicon contact does not affect collector current but reduces base current, which is, however, also influenced by transport effects in the monocrystalline part of the emitter. To extract quantitative information of the polysilicon contact the hole current reaching the poly/mono-Si interface

$$j_p = qS\,(P_n - P_{no})$$

is described by the effective recombination velocity S /1,2/. We have analyzed measured base currents by solving the transport equations in the mono crystalline part of the emitter with the device simulator MEDUSA /3/ for different boundary conditions S. Fig. 1 shows calculated and measured base currents as a function of emitter junction depth. In the case of the metal contacted emitter ($S = \infty$) the base current increases with decreasing junction depth indicating that the emitter becomes more and more transparent to hole transport. In contrast the "ideal" poly-Si contact ($S = 0$) imposes that recombination only occurs in the mono crystalline part of the emitter resulting in the low base currents at small emitter junction depths. The intermediate curve ($S = 7 \cdot 10^4$ cm/sec) is obtained for a poly-emitter contact given an HF-dip as interface treatment. For this type of device the base current is nearly independent on emitter junction depth.

Surface recombination velocity and thus I_{bo} depends on interface treatment as well as on annealing temperature. We have fabricated devices annealed at various temperatures (850°C - 950°C) and times. To minimize oxide contamination of the poly/mono-Si interface all devices were given an HF-dip as an interface treatment. The base saturation currents and specific emitter-resistances are shown in fig. 2 as a function of x_{je}. Although a significant reduction in S and thus I_{bo} is found by low temperature processing (and thus low x_{je}) this direction of progress is limited by a corresponding increase in emitter resistance. These results can be correlated with HXTEM-micrographs of the interface. At low annealing temperatures a continuous interfacial oxide layer with a thickness varying from 8 Å to 12 Å is found. At higher annealing temperature a break up of the interfacial oxide layer occurs.

The increase of emitter resistance at very shallow emitter junctions can be circumvented by rapid optical annealing (ROA). ROA is more effective in breaking up interfacial oxide layers resulting in emitter resistances of about 30 $\Omega\mu m^2$ at emitter junction depths of about 50 nm /4/.

The larger current gain and emitter resistance that occurs in the case of a continuous oxide film can be explained in terms of tunneling /5,6/. The larger bandgap of the insulating film forms a barrier both to electrons and holes. Assuming an effective barrier height of $E_n = 0.3$ eV, S can be calculated as a function of oxide layer thickness. The results are shown in fig. 3 with and without an additionally assumed recombination at surface traps. In the case of continuous interfacial oxide layers the tunneling model fits the "experimental" results - as obtained from measured base currents and HXTEM-studies - quite well.

3. BASE CHARGE CONTROL

A major driving force for scaling down emitter depths, of course, is the goal to achieve extremely narrow base widths W_B. Several problems are encountered with conventional boron ion implantation if W_B values below approximately 200 nm are desired. Due to the strong channeling effect for boron ions, either very low implantation energies (<10 keV) have to be chosen or a thick screen oxide (>100 nm) has to be used for the active base implant. In either case control of base charge (and thus current gain) is extremely difficult, essentially because the peak of the boron distribution lies very close to or even above the silicon surface. Base charge, therefore, is highly sensitive to variations in emitter junction depth.

As can be seen from fig. 4, achieving W_B=120 nm and requiring x_{jE} to be not larger than the peak boron penetration depth x_p (without screen oxide) would already lead to x_{jE} as low as 30 nm.

New methods for active base doping have to be investigated therefore. A very attractive one is double diffusion of boron and arsenic out of the emitter poly-Si layer, because in this case the reference plane for both diffusion profiles is the same poly-Si/mono-Si interface. Unfortunately, both diffusions can not be performed within a single heat cycle, because boron is nearly immobile in the presence of high concentrations of arsenic (fig. 5a). First diffusing boron out of the poly-Si, then implanting and diffusing arsenic yields very narrow base widths (fig. 5b) and has already resulted in very high f_t devices /7/. However, base charge control again is very difficult because the slope of the boron profile at x_{jE} is very high. This sets extremely stringent requirements for temperature control of the two diffusion steps and for poly/mono-interface homogeneity.

Preamorphization of the single crystal silicon with, e.g., a germanium implant before implanting the active base can be used to avoid boron channeling. The problem with this method, of course, is defect generation. This issue is therefore presently investigated in detail. Preliminary results show in agreement with /8,9/ that for amorphization depths of 50 nm and below perfect annealing of the implant damage should be possible.

4. TRANSIT TIME CONSIDERATIONS

One of the most important factors determining speed of bipolar circuits is the transit time τ_F, which in first order calculations is proportional to W_B^2. To take into account second order effects like carrier mobility and electrical field in the base, we have determined τ_F for various base doping profiles numerically, assuming Gaussion distributions both for emitter and base profile. Carrier distributions were calculated using the program MEDUSA, then the carrier densities were integrated over the whole transistor for different applied voltages, diffusion charges were extracted, and $\tau_F = dQ_{diff}/dI_C$ was calculated.

As expected, τ_F decreases more than linearly with decreasing W_B (fig. 6). On the other hand, if one increases the doping concentration N_A in order to achieve a low base sheet resistance R_{pinch}, τ_F increases. This effect is due to the concentration dependent mobility in the base and to charge storage in the emitter.

There are some constraints in choosing the optimum base width W_B and base doping N_A with respect to transit time τ_F. First, reducing W_B keeping N_A constant leads to an increasing R_{pinch}.

It can be shown, that for emitter stripe widths as small as 0.5 µm the logic gate delay degradation due to increasing base pinch resistance is small compared to the gain due to transit time reduction. However, punch-through must be avoided at normal operating voltages. The dashed line in fig. 6 indicates punch-through occuring at U_{BC} = 3V and U_{BE} = 0V. Depending on N_A, this condition corresponds to R_{pinch} values (at U_{BE} = 0V) from about 30 kΩ to about 60 kΩ.

A second constraint in optimizing the base doping profile is demonstrated in fig. 7. Here we have plotted measured base current characteristics for transistors with base doping concentrations N_A up to $3 \cdot 10^{19}$ cm^{-3}. At concentrations exceeding $N_A = 1 \cdot 10^{19}$ cm^{-3} the base current is severely degraded due to forward tunneling in the emitter base junction. Taking this value of $1 \cdot 10^{19}$ cm^{-3} as an upper limit for the base doping concentration and requiring that no punch-through occurs at U_{BC} = 3V, a transit time of approximately 4 ps could be achieved at W_B = 20 nm.

We now investigate the contribution of emitter charge storage to forward transit time. In our first example we assume W_B = 50 nm, x_{jE} = 40 nm, S = $7 \cdot 10^4$ cm/sec and $N_D^+ = 7 \cdot 10^{19}$ cm^{-3}. The results for emitter and base transit time are plotted in fig. 8 as a function of base doping N_A. Due to the reduction in electron mobility base transit time first increases with N_A. At higher base dopings electron mobility saturates and the drift current, caused by the internal field, increases resulting in nearly constant values of τ_B. In contrast τ_E increases with $1/I_C$ and at doping levels $N_A > 5 \cdot 10^{18}$ cm^{-3} the change in total transit time is mainly due to an increase in τ_E. For typical current gains β = 100 we find τ_E = 1.7 ps which has to compared with a base transit time of 2.3 ps. In fig. 9 τ_E is plotted as a function of emitter depth for various values of S and electrical activation N_D^+. The base profile (W_B = 50 nm, $N_A = 5 \cdot 10^{18}$ cm^{-3}) and impurity gradient of the emitter was kept fixed during the calculations.

From our results in figs. 8 and 9 we conclude that for optimum performance of downscaled transistors minimization of both τ_E and τ_B is of importance. To decrease minority carrier storage in the emitter as far as possible very shallow junctions and a high electrical activation is needed.

5. CONCLUSIONS

Improvements in transistor current gain by using polysilicon emitters are limited by the detrimental influence of interfacial oxides on emitter resistance, which is adequately described by a tunneling model. Taking this constraint into account, a polysilicon emitter does not significantly improve current gain with respect to an opaque emitter. It rather allows one to reduce x_{jE} to virtually zero without significantly changing β, thus giving additional freedom in process and device design.

Using double diffusion techniques or preamorphization implants, base widths below 50 nm should thus be obtainable, although serious problems concerning process control and defect generation still have to be solved. A lower limit to the achievable basewidth is, on the other hand, set by emitter-collector punch-through and base doping limitations due to tunneling.

From these considerations, a lower limit of around 4 ps is expected for the forward transit time τ_F of pure silicon bipolar transistors, which corresponds to f_T around 40 GHz. The contribution of emitter charge storage to τ_F becomes increasingly important for these shallow devices and may amount to about 30 % at 50 nm base width. This again stresses the need for extremely shallow, but highly doped, emitters as provided by the polysilicon emitter technology.

A possibility to further improve intrinsic device speed over the figures given above would be the use of true heterojunction emitters which, among other advantages, totally eleminate emitter charge storage.

6. REFERENCES

/1/ J.G. Fossum, M.A. Shibib, IEDM Techn. Dig. (1980) 280

/2/ B. Benna, T.F. Meister, H. Schaber, A. Wieder, IEDM Techn. Dig. (1985) 302

/3/ W.L. Engl: "MEDUSA User Manual"; ITHE Aachen; Techn. Hochschule Aachen (1985)

/4/ H.J. Böhm, H. Kabza, T.F. Meister, H. Wendt; Spring Meeting Electrochem. Soc., Philadelphia, May 1987, No. 187

/5/ Z. Yu, B. Ricco, R.W. Dutton; IEEE Trans. Electron Dev.; (1984) 773

/6/ B. Benna, T.F. Meister, H. Schaber; to be published in Solid-State Electr.

/7/ H.K. Park, K. Boyer, A. Tang, C. Clawson, S. Yu, T. Yamaguchi, J. Sachitano; Proc. 1986 Bipolar Circuits and Technology Meeting, Minneapolis, September 1986, p. 39

/8/ A.C. Ajmera, G.A. Rozgonyi; Extended Abstr. Spring Meeting Electrochem. Soc. Boston, May 1986, No. 239

/9/ M.C. Ozturk, C. Lee, J.J. Wortman; Spring Meeting Electrochem. Soc., Philadelphia, May 1987, No. 235

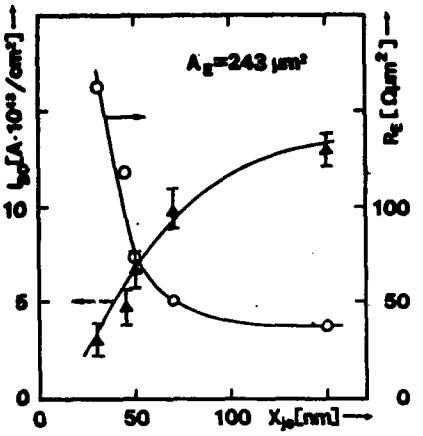


Fig. 2: Base saturation current and specific emitter contact resistance for a range of emitter drive in temperatures.

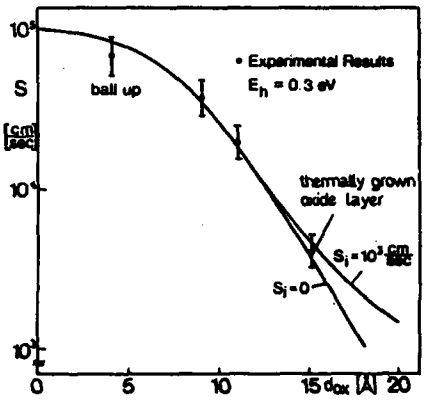


Fig. 3: Effective recombination velocities as a function of oxide layer thickness.

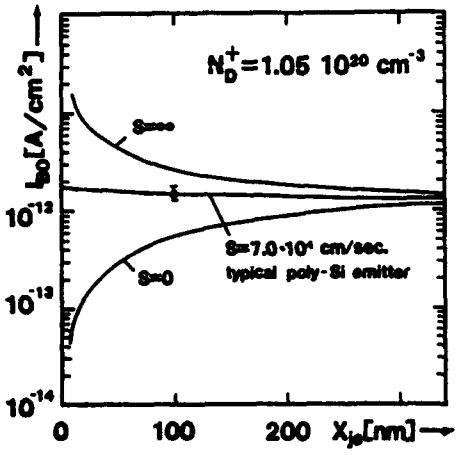


Fig. 1: Base saturation current as a function of emitter junction depth for different interface conditions.

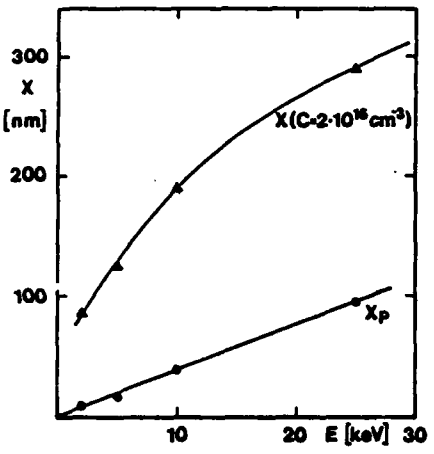


Fig. 4: Peak boron penetration depth x_p and base-collector junction depth x (at $2 \cdot 10^{18}$ cm^{-3}) obtained from SIMS measurements.

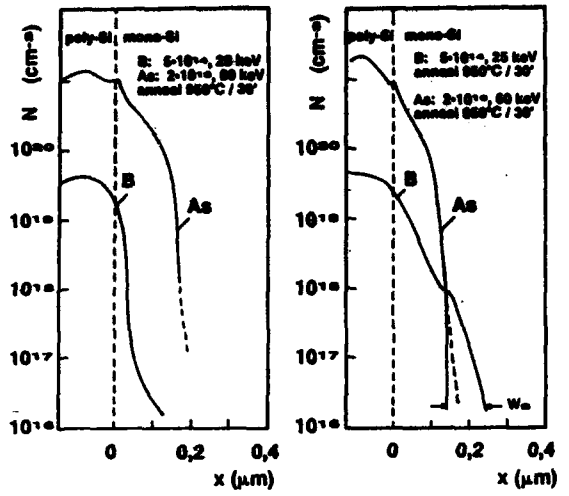


Fig. 5: Results of joint (a) and successive (b) diffusion of boron and arsenic from a single poly-Si layer.

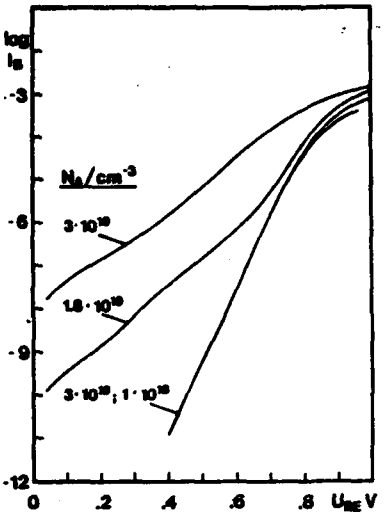


Fig. 7: Base current characteristics for transistors with different base doping concentrations.

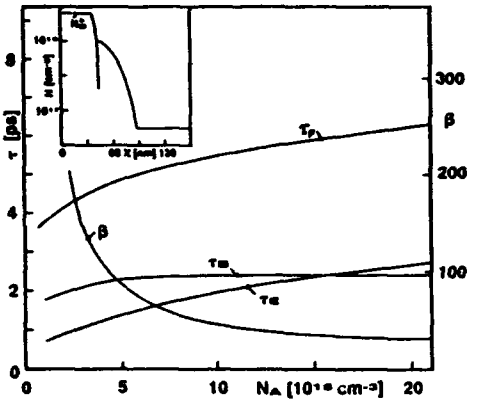


Fig. 8: Forward transit times τ_E, τ_B and τ_F as a function of base doping level N_A. The base width (W_B = 50 nm) and emitter profile ($N_D^+ = 7\cdot10^{20}$ cm^{-3}, x_{jE} = 40 nm, S = $7\cdot10^4$ cm/sec) were kept fixed during the calculations.

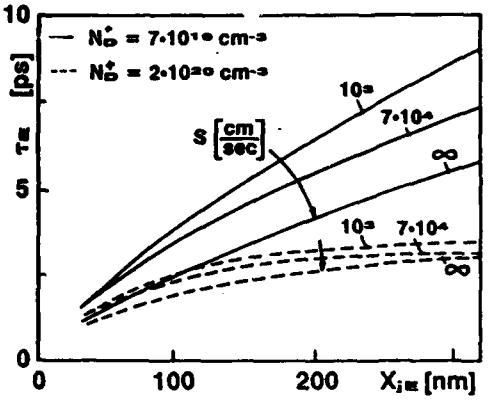


Fig. 9: Emitter transit time τ_E as a function of emitter junction depth for various values of electrical activation and surface recombination velocity S.

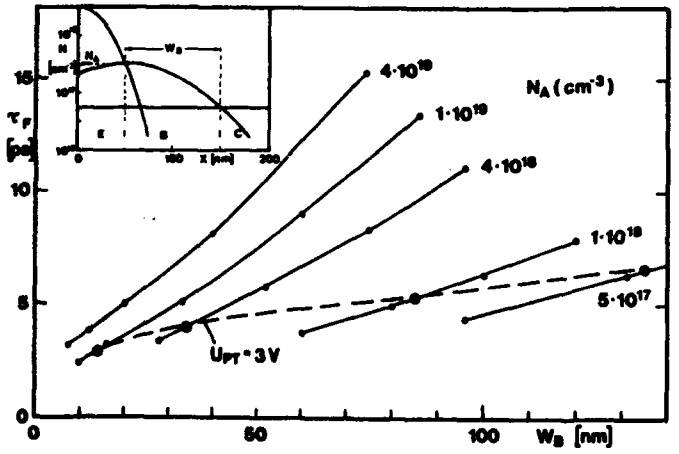


Fig. 6:
Transit time τ_F vs. base width W_B for different base doping N_A. Dashed line: punch-through at U_{BC} = 3V.

TRENCH ISOLATION SCHEMES FOR BIPOLAR DEVICES --- BENEFITS AND LIMITING ASPECTS

Hiroshi GOTO and Katsuyuki INAYOSHI

Bipolar Process Division, Fujitsu Limited
1015, Kamikodanaka, Nakahara, Kawasaki, 211, Japan

This paper gives a review of benefits and limiting aspects in trench isolation techniques for bipolar devices. The most sophisticated trench isolation techniques have realized not only higher packing densities but reduced collector-substrate, wiring-substrate and base-collector parasitic capacitances. By using these techniques, high performance bipolar devices have been fabricated while crystal defects caused by trench structures are the serious problem. Trench isolaltion techniques are still in progress, and it seems now that there is no apparent limiting aspect until the trench width reaches the filler material width to sustain enough breakdown voltage.

1. INTRODUCTION

Since the first application of trench isolation to 1Kb ECL RAMs in 1982 [1], a lot of bipolar devices have been fabricated by various kinds of trench isolation techniques. In the case of bipolar devices, some trench isolation techniques have been used in volume production since the early stage of the development, though trench isolation in MOS devices is in just experimental trials. This is partly because isolation in bipolar devices plays more important role in their performances than MOS devices.

In the following sections, process technologies, device structures, benefits and limiting aspects in trench isolation for bipolar devices will be discussed.

2. PROCESS TECHNOLOGIES AND DEVICE STRUCTURES

Trench isolation was first disclosed in 1978 [2], and its application to devices appeared in early 1980's [1], [3]. These trench isolation techniques consisted of three process steps; trench etching, trench filling and planarization. The typical structure of narrow, deep trenches for isolation is achieved by anisotropic plasma etching such as RIE. After trenches are etched, their surfaces should be covered by insulators in order to accomplish

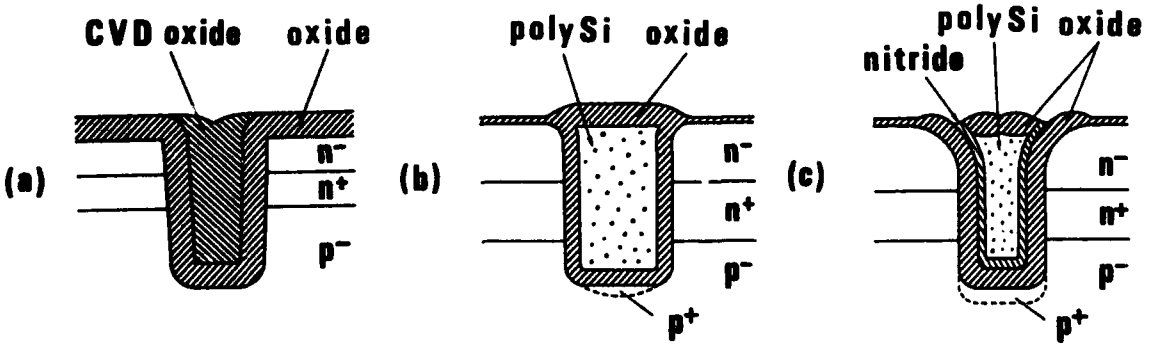


Figure 1 Typical trench isolated structures. (a)DGI [2], (b)IOP-II [1], (c)U-Iso [3]

electrical isolation. Thermal oxide is commonly used to coat the trench surfaces. Then trenches are filled with dielectric materials such as CVD oxide, nitride, undoped polysilicon and so on. Not only trench etching but this filling and following planarization steps are critical because unless the trenches are completely filled with these dielectric materials, the voids in trenches cause sharp crevasses in isolation regions. Uniform planarizing techniques are strongly required to attain planar surfaces, otherwise sharp steps occur at the trench edges. These crevasses and uneven steps make metallization step coverage unfavorable. For planarizing, controlled etch-back techniques or polishing techniques are used. Figure 1 shows three typical trench isolated structures which may be called the first generation.

After the first generation, some structural modifications for trench isolation followed. These modifications were mainly focused on the reduction of wiring-substrate or base-collector parasitic capacitances. In the case of previous structures, only the isolation regions were reduced, and other inactive regions such as field regions for wiring were left without any thick insulator. Base regions were not optimally isolated from collector reach-through regions, either. By combining LOCOS or shallow moats or trenches with deep trenches, these disadvantages of the first generation have been overcome. Figure 2 shows those modified trench isolated structures. Figure 2 (a) is an example of CVD oxide-filled trenches combined with ROX (Recessed OXide) [4], and Figure 2 (b) shows the dual-depth trenches filled with oxide, nitride and undoped polysilicon [5]. Shallower trenches are used for SBD-base or base-collector isolation. Figure 2 (c) is a U-FOX (U-groove isolation with thick Field OXide; renamed from IOP-L [6]) structure. In the case of U-FOX, LOCOS step is performed prior to the trench isolation step, and a planar surface is achieved. Figure 2 (d) shows deep trenches and shallow moats filled with oxide [7]. By optimizing the distance between deep trenches in shallow moats, void formation is elaborately eliminated. Figure 2 (e) is a modification of Figure 2 (a)[8]. CVD oxide is replaced by selective-epi-silicon to avoid voids in trenches

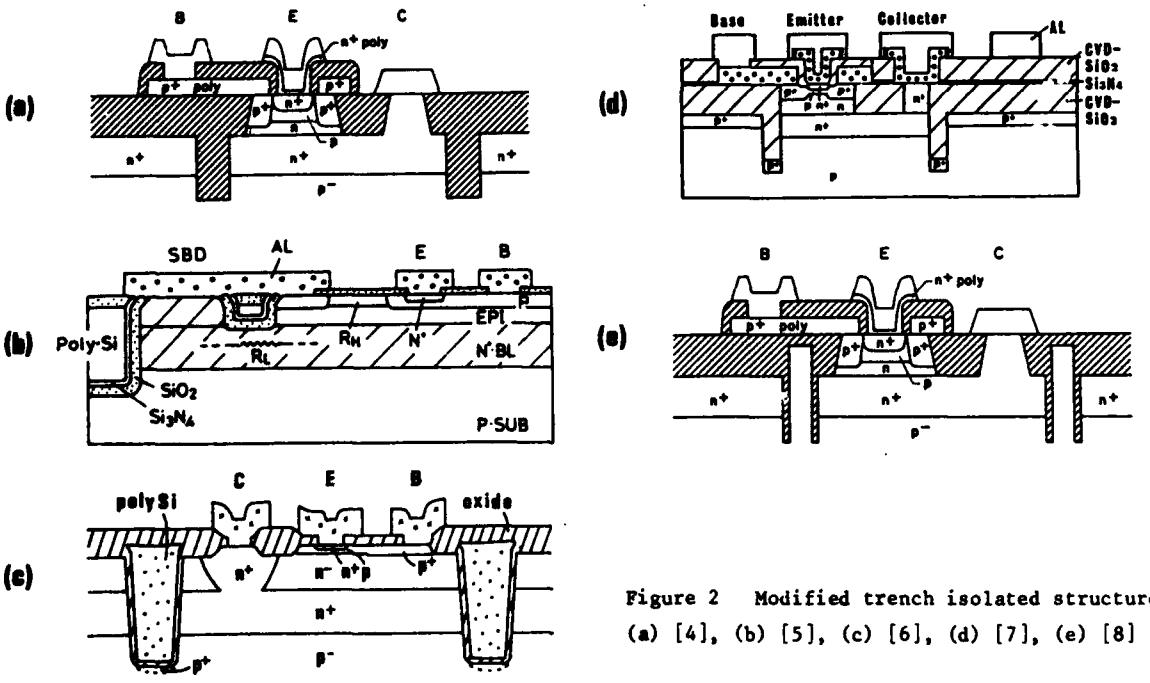


Figure 2 Modified trench isolated structures. (a) [4], (b) [5], (c) [6], (d) [7], (e) [8]

Table 1 Bipolar devices using trench isolation techniques (excluding ring oscillators).

Laboratory	Device	Performance	Process	Reference
Fujitsu	1Kb ECL RAM	Taa=4.4ns	IOP-II; (*)	IEDM 1982
	4Kb ECL RAM	Taa=3.5ns		ISSCC 1983
	16Kb ECL RAM	Taa=15ns		ISSCC 1983
	64Kb ECL RAM	Taa=10ns		ISSCC 1985
	16Kb ECL RAM & 1.2K Gate Array	Taa=2.8ns & Tpd=280ps		ISSCC 1986
	64Kb ECL RAM	Taa=5ns		ISSCC 1987
	Prescaler	Ft =1.6GHz	U-FOX; LOCOS & IOP-II	IEDM 1984
Hitachi	4Kb ECL RAM	Taa=2.5ns	U-groove; (**)	VLSI Symp. 1984
	16Kb ECL RAM	Taa=3.5ns		ISSCC 1986
NEC	16Kb ECL RAM	Taa=4ns	Trench; (no detail)	ISSCC 1986
NTT	ECL 5K G.A.	Tpd=165ps	SST & Trench; (*)	CSSDM, Tokyo 1985
IBM	32Kb RAM	Taa=3ns	Polysilicon-filled Trench	ISSCC 1986
	5Kb ECL RAM	Taa=1.0ns	ROX & Selective-epi-silicon -filled Trench	ICSSDM, Tokyo 1986
	32b Processor	Tc =60ns	Deep Trench; (no detail)	ISSCC 1986
AMD	128Kb PROM	Taa=35ns	IMOX-III-Slot; (*)	ISSCC 1986
Fairchild	16x4 Reg. File	Taa=0.9ns	SAPT; (*)	ISSCC 1986
Honeywell	PLA	Tpd x Power =26fJ	ADB-III; (Trench filled with oxide)	VLSI DESIGN Jan. 1985

Taa; address access time Tpd; basic gate delay Ft; toggle frequency Tc; cycle time

(*) ; Trench filled with oxide and polysilicon
(**); Dual-depth trench filled with oxide, nitride and polysilicon

and to omit the planarizing step.

3. BENEFITS AND APPLICATION TO DEVICES

The sophisticated trench isolation techniques have following advantages.

1) Higher packing densities than conventional oxide isolation.
2) Reduction of collector-substrate parasitic capacitances.
3) By combining LOCOS or shallow trenches or moats, reduction of wiring-substrate and base-collector parasitic capacitances.

Accordingly, high performance bipolar devices have been demonstrated by using them. Table 1 shows the list of integrated circuits except ring oscillators made by trench isolation techniques reported to date. In 1983 only two devices were reported at ISSCC, while in 1986 there were seven devices of not only memories but logic and microprocessors.

4. PROBLEMS AND LIMITING ASPECTS

Trench isolation is now getting essential for bipolar devices and several companies are manufacturing devices in volume production. But we cannot say that trench isolation is matured to its perfect status. Some problems are still left unsolved. These problems are mainly divided into two categories. One is the structural problems, and the other is the electrical characteristics ones. They are often strongly related and the former is apt to influence the latter. If the trench etching process is not optimized, the trench shapes change variously. When the trench sidewall becomes barrel-like shape or an overhanged structure, the filling material cannot be filled inside the trenches completely. After the controlled etch-back

process, voids appear in the isolation regions. These cause the unfavorable metallization step coverage. Lateral etching of highly doped buried layers, silicon needles called black silicon and trenching at bottom edges also make trench shapes maladapted. When the surface of these trenches is oxidized, strong stresses occur at the sharp curved edges and cause crystal defects such as dislocation. They are often major reasons of electrical leakage. The most serious problem is how to get rid of the crystal defects. They seem to be strongly related to concentrated stresses, but the mechanism is not yet clear although some process modifications have been tried [9], [10].

Even if there are some problems mentioned above, trench isolation techniques are in progress, and it seems there is no limiting aspect in the future. At present, trench widths are limited by lithographic systems, not by other factors such as etching or deposition. However, the widths cannot be reduced less than the coating or filler material widths. So, the thickness of insulator needed for satisfied electrical isolation limits the trench width. Air isolation is the ultimate goal of trench isolation [11]. But before reaching there, other process obstacles such as metallization reliability should be removed and they are the limiting factors for scaling down.

5. CONCLUSION

Trench isolation is now one of the key technologies in high performance bipolar devices. It realizes higher packing densities and reduced collector-substrate capacitances. It also reduces wiring-substrate and base-collector capacitances with process modifications.

Trench isolation, as well as so called sophisticated self-aligned structures such as SST or SICOS, is leading the bipolar technology towards the ultra large scale and high performance integrated circuits.

REFERENCES

[1] Goto, H., Takada, T., Abe, R., Kawabe, Y., Oami, K. and Tanaka, M., IEDM Tech. Dig. (1982) 58

[2] Bonder, J.A. and Pogge.H.B., U.S.Patent 4104086 (1978)

[3] Hayasaka, A., Tamaki, Y., Kawamura, M., Ogiue, K. and Ohwaki, S., IEDM Tech. Dig. (1982) 62

[4] Tang, D.D., Solomon, P.M., Ning, T. H., Isaac, R.D. and Burger, R.E., ISSCC Tech. Dig. (1982) 242

[5] Tamaki, Y., Shiba, T., Homma, N., Mizuno, S. and Hayasaka, A., Symposium on VLSI Technology Tech. Dig. (1983) 24

[6] Goto, H., Takada, T., Nawata, K. and Kanai, Y., Symposium on VLSI Technology Tech. Dig. (1985) 42

[7] Sakai, H., Kikuchi, K., Kameyama, S., Kajiyama, M. and Komeda, T., Symposium on VLSI Technology Tech. Dig. (1987) 17

[8] Tang, D.D., Li, G.P., Chuang, C.T., Danner, D., Ketchen, M.B., Mauer, J., Smyth, M., Manny, M., Crossler, J.D., Ginsberg, B., Petrillo, E. and Ning, T.H., ISSCC Tech. Dig. (1986) 104

[9] Teng, C.W., Slawinski, C. and Hunter, W.R., IEDM Tech. Dig. (1984) 586

[10] Sagara, K., Tamaki, Y. and Kawamura, M.,J. Electrochem. Scc., vol. 134 (Feb. 1987) 500

[11] Riseman, J., U.S.Patent 4106050 (1978)

A2.2.4

A SALICIDE BASE CONTACT TECHNOLOGY (SCOT) FOR USE IN HIGH SPEED BIPOLAR VLSI

Tadashi HIRAO, Tatsuhiko IKEDA and Yoichi KURAMISU

LSI Research and Development Laboratory,
Mitsubishi Electric Corporation,
4-1 Mizuhara, Itami, P.O.Box. 664, Japan.

This paper describes a new process technology, which is called SCOT : salicide (self-aligned silicide) base contact technology, and applied for realizing high performance prescaler IC and high gate density masterslice LSI. The main feature of this process, for reduction of the base resistance and capacitance, is a silicidation of the base contact which is opened by employing self-alignment technology. A 1/128, 1/129 two-modulus prescaler IC constructed of the 1.5 μm SCOT transistors has been improved to a high operation of 2.1 GHz at 56-mW power dissipation. An ECL 18K-gate masterslice has been developed by a variable size cell (VSC) approach, employing the SCOT process.

1. INTRODUCTION

The two-modulus prescaler ICs used for automobile telephones and satellite communication receivers have been required not only to operate at a high frequency such as about the GHz band but also to operate with the low power dissipation. The 1.6-GHz 63-mW Si prescaler IC (1) and 1.8-GHz 46-mW GaAs prescaler IC (2) have been reported. In order to enhance the performance of the prescaler ICs, it is important to set up an optimum transistor for the high frequency operate at the low power dissipation. A new proposed transistor was realized with the base contact of a salicide structure which was made by silicidation of poly-Si and silicon surface simultaneously in the same way as MOS technology (3). Using this salicide base contact technology (SCOT) process (4), the transistor characteristics were improved and then the high performance of two-modulus prescaler ICs was obtained.

For high speed data processing systems such as computer mainframes, the ECL masterslice LSIs have been required to increase the integration degree as well as the operating speed. Several types of ECL masterslices (5),(6) have been developed as suitable to these requirements. A relatively large number of the nonutilized elements, however, remain in these masterslices. A variable size cell (VSC) approach (7) employing the SCOT process is proposed to reduce the nonutilized transistors and resistors.

2. PROCESS AND TRANSISTOR DESIGN

Fig.1(A) illustrates the top view of SCOT transistor and Fig.2 shows the key step in the process sequence. The isolation of SCOT transistor comprised the full-recessed oxide and the semi-recessed oxide surrounded on the base area, as shown in Fig.2(A). The P^+ poly-Si layer was formed on the base edge exteded over the



FIGURE 1

Schematic layout of a SCOT transistor and ISAC transistor.

the isolation oxide and the poly-Si resistor was fabricated simultaneously. After the base formation, the emitter region was produced by the diffusion from the implanted N^+ poly-Si. The oxide on the base contact and P^+ poly-Si was etched away by using the photo-resist mask of emitter poly-Si etching as shown in Fig.2(C). The thick oxide was selectively grown over the heavily arsenic doped emitter poly-Si. Therfore, the oxide covering the emitter poly-Si remained after removing the thin oxide on the base contact and P^+ poly-Si, and separated the base contact from the N^+ poly-Si. The Pt-silicide was formed both on the epi-surface and P^+ poly-Si of base contact, as shown in Fig.2(D). The base electrode was fabricated with both the silicide of self-align opened base contact and the polycide, and then this can be called a salicide base contact. Opening the contact windows and Al-metallization completed the processing of SCOT transistor, as shown in Fig.2(E).

As a result of the gate speed analysis of the transistor characteristics (4), the important parameters for high speed performance are not only the cuttoff frequency f_T, the collector-base capacitance C_{TC} but also the base series resistance r_B . It was learned by the following

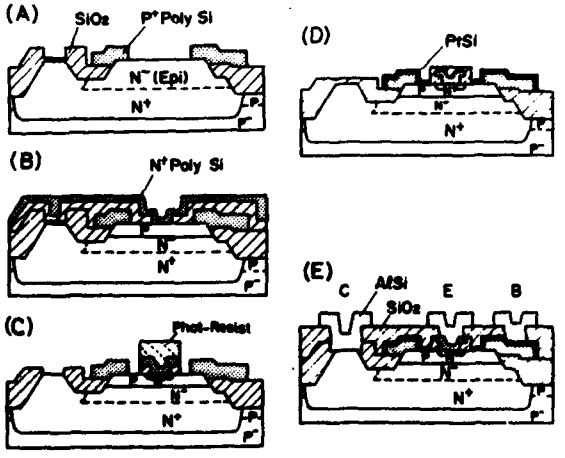


FIGURE 2

Fabrication procedure of SCOT transistor.

comparison with the coventional ISAC (8) transistor shown in Fig.1(B) that the SCOT transistor has realized the optimum transistor design with reduction of these characteristics. The r_B reduction in SCOT transistor was done by decreasing the distance D between the base contact and emitter and by the double base structure. The distance D , in the case of SCOT transistor, is determined by the salicide base contact structure and is close to 1 μm, a half value of that in ISAC one. The base area in SCOT transistor is decreased to about one-half of that in ISAC one. In order to decrease the base area, the emitter length can be decreased as the goal to maintain a small r_B. The reduction of parasitic base region, furthermore, was achieved by the full walled base structure with the semi-recessed oxide and by the salicide base contact structure. The vertical down-scaling, in which the emitter depth is 0.1 μm and the base width is 0.13 μm, produces the higher f_T. Table 1 shows that a SCOT transistor has been synthesized in the transistor design for high performance.

3. GATE SPEED AND PRESCALER IC

The performance of SCOT transistor used for a prescaler IC was improved as compared with the current product with ISAC process, that is,the C_{TC} and r_B decreased to half value and f_T became twice as high, as shown in Table 1. The maximum f_T obtained, moreover, is 9.5 GHz at I_C = 6 mA by the SCOT transistor in which the emitter is

TABLE 1

Comparison of features of newly developed prescaler and current prescaler.

		New development	Current product
Process technology		SCOT	ISAC
Emitter size		1.5 x 3 μm^2	1.5 x 5 μm^2
Emitter depth		0.1 μm	0.4 μm
Capacitance	C_{TC}	9 fF	20 fF
Base resistance	r_B	49 Ω	92 Ω
Cutoff frequency	f_T	4.1 GHz	2.2 GHz
Delay time of Ring-osc.		140 ps	267 ps
Max. operating frequency		2.1 GHz	1.1 GHz
Power dissipation		56 mW	125 mW

four-fingers of 1.5 x 5 μm^2.

It is demonstrated in Fig.3 that the gate speed t_{pd} of the ECL ring-oscillator employing SCOT transistor on same emitter size (1.5 x 5 μm^2) was faster than that of ISAC one, especially at the high current range by effect of reducing r_B. The mimimum t_{pd} achieved 116 ps at gate current I_g = 1.3 mA. Simulating t_{pd} in scaling down of the SCOT transistor to 1.0 μm emitter, furthermore, the high speed can been realized at less than 80 ps.

A 1/128,1/129 two-modulus prescaler IC was fabricated with 1.5 μm SCOT transistors and poly-Si resistors. First level metallization of AlSi and polyside cross-under interconnection were employed. This prescaler IC operated in a wide range from 400 MHz to 2.1 GHz with 56 mW at 5-V supply voltage. In Fig.4, the performance of SCOT prescaler IC was compared with other products. The SCOT prescaler IC had four times higher performance than the ISAC prescaler. This prescaler IC operated with the half power dissipation of the reported Si prescaler (1) and with the same one of the GaAs prescaler (2). Decreasing the power dissipation, we obtained 1.4 GHZ with 30 mW and 850 MHz with only 19 mW.

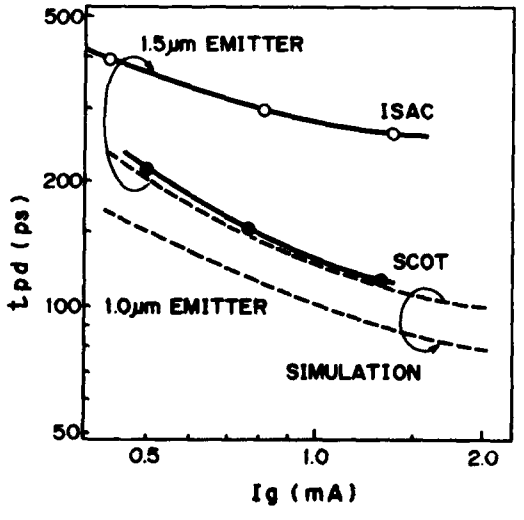


FIGURE 3

Relations between gate speed and gate current.

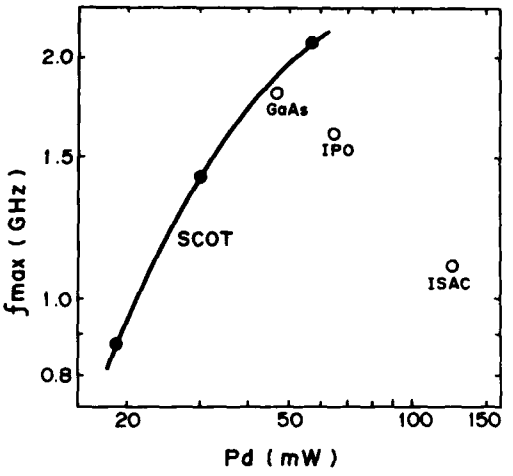


FIGURE 4

Comparison of maximum operating frequency and power dissipation.

4. VSC MASTERSLICE

The actual pattern layout of two-input OR/NOR gate is shown in Fig.5. In the case of the current cell approach, a relatively large number of the nonutilized elements ramain, for example the macrocell array MCA in the simple logic functions. The VSC concept is based on the design of an array which is constructed from cellular units. This VSC construction was found to reduce the nonutilized elements.

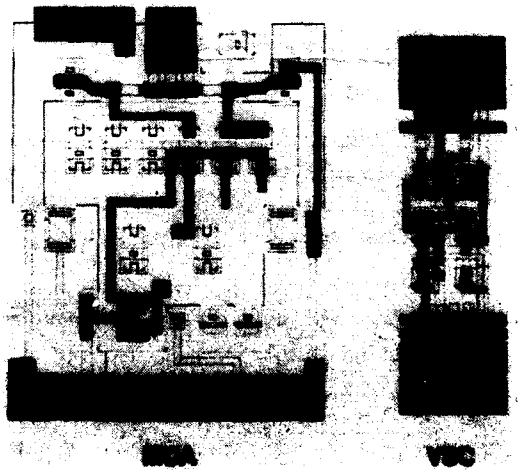


FIGURE 5

Layout pattern of two-input OR/NOR gate.

Fig.6 shows a schematic diagram of SCOT transistor and polycide interconnection used for VSC masterslice. To realize a VSC structure, the proper use of poly-Si patterns was required during the slice process. The resistor value in each logic cell was determined by the silicidation of poly-Si pattern. In addition, unused poly-Si patterns can be utilized for the policide interconnection.

An ECL 18K-gate masterslice (7) was developed by VSC approach and fabricated by employing 1.5 μm SCOT process with four-level metallization. The features of the VSC masterslice are summarized in Table 2. The gate density was increased by more than 20 % in the VSC structure compared with the current cell structures. The basic gate delay of 150 ps was attained at the power dissipation of 2.4 mW.

5. CONCLUSION

As the excellent structure with reduction of C_{TC}, f_T and r_B simultaneously, the SCOT process by employing self-alignment silicide technology has been proposed. A two-modulus prescaler IC has achieved 2.1-GHz operation with 56-mW power dissipation. An ECL 18K-gate masterslice has been developed by a VSC approch which maximized the utilization of elements.

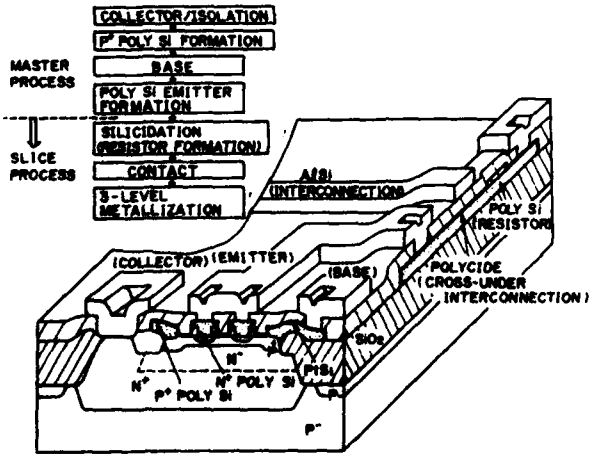


FIGURE 6

Schematic diagram of SCOT transistor and polycide interconnection used for VSC masterslice.

TABLE 2

Features of VSC masterslice.

Technology	1.5 μm rule SCOT
No. of transistors	39,936
No. of poly-Si resistors	53,248
No. of units	13,312
Unit size	24 μm X 204 μm
Metal pitch	1st 8 μm
	2nd 6 μm
	3rd 8 μm
No. of channels	1,936
No. of I/O pins	256
Interface	ECL 100K compatible
Intrinsic gate delay	150 ps
Supply voltage	V_{EE} : -4.5 V
	V_{TT} : -2.0 V
Switching current	0.4 mA
Emitter-follower current	0.3 mA/0.6 mA
Chip size	11.90 mm X 11.96 mm

ACKNOWLEGEMENTS

The author would like to thank N.Katoh and T.Nishimura for circuit deign, and K.Sakaue and Y.Kinosita for wafer processing. We would also like to thank N.Tsubochi for fruitful discussions and express gratitude to K.Shibayama and H.Nakata for their encouragement.

REFERENCES

(1) H.Suzuki, T.Akiyama and K.Ueno, IEDM Tech. Digest (1984) 682

(2) K.Maemura, T.Takahashi, S.Inoue, Y.Mitsui, S.Orisaka, O.Ishihara and M.Otsubo, IEDM Tech. Digest (1985) 94

(3) C.Y.Ting, IEDM Tech. Digest (1984) 110

(4) T.Hirao, T.Ikeda and N.Katho, Extended Abstr. of 17th Conf. SSDM (1985) 381

(5) W.Brackelmann, H.Fritzsche, H.Ullrich and A. Wieder, IEEE J. Solid-State Circuits, SC-20 (1985) 1032

(6) M.Tatsuki, S.Kato, M.Okabe, H.Yakushiji and Y.Kuramistu, IEEE J. Solid-State Circuits, SC-21 (1986) 234

(7) T.Nishimura, H.Sato, M.Tatsuki, T.Hirao and Y.Kuramitsu, IEEE J. Solid-State Circuits, SC-21 (1986) 727

(8) Y.Akasaka, Y.Tsukamoto, T.Sakurai, T.Hirao, Y.Horiba, K.Kijima and H.Nakata, IEDM Tech. Digest (1978) 189

TRENDS IN HETEROJUNCTION SILICON BIPOLAR TRANSISTORS

R. MERTENS, J. NIJS, J. SYMONS, K. BAERT and M. GHANNAM

Interuniversity Microelectronics Center (IMEC)
Kapeldreef 75
B-3030 Leuven, Belgium

Different types of bipolar transistor emitters are described. Epitaxial emitters can be achieved by solid phase epitaxial regrowth of polysilicon (at T > 850°C) and recently by glow discharge deposition at T = 250°C and recrystallization (at T = 700°C). Wide band gap emitters and narrow bandgap bases result in very high emitter efficiency which has to be traded-off with emitter and base series resistances.

1. INTRODUCTION

VLSI bipolar transistors typically have polysilicon emitters to achieve high emitter efficiency and large packing density. In recent years [1,2] it has become clear that for optimum performance of polysilicon emitters in VLSI applications a trade-off must be made between emitter efficiency and emitter series resistance. Work at several laboratories has indeed clearly shown that the emitter Gummel number (GN_E) is not the appropriate figure of merit of a modern emitter-base junction used in high speed VLSI applications. The true figure of merit depends on GN_E and also on the emitter series resistance.

In this paper different emitter-base structures aiming at high emitter efficiency and low emitter series resistance will be discussed. In a first part of the paper our work on epitaxial emitters will be described. Although epitaxial emitters are not heterojunctions in the strict sense of the word, they are included here because epitaxy is a technology closely related to heterojunction processing and since in the short term these emitter may turn out to be the best alternative for poly-emitters.

The second part of the paper will deal with true heterotype emitters starting with an overview of the different wide-gap emitters that have been proposed in the literature. Finally, the paper will end with a discussion about the possibilities of narrow base transistors.

2. SILICON BIPOLAR TRANSISTORS WITH EPITAXIAL EMITTERS

As pointed out in the introduction polysilicon emitters are suffering from a trade-off that has to be made between emitter efficiency and emitter series resistance. Recent work [1] has indicated that high emitter efficiency only can be realized at the expense of a large series resistance. Such a specific series resistance R_E (ohm.cm^2) is detrimental for high speed performance if [3]:

$$R_E > \frac{kT}{J} \qquad (1)$$

In (1) J is the emitter current density which in advanced bipolar transistors can be as large as 10^5 A/cm^2. Formula (1) predicts that R_E should be lower than 2.6×10^{-7} A/cm^2 to eliminate degradation of the transconductance. This is a very low value which is difficult to achieve with polysilicon emitters. The emitter series resistance of a polysilicon emitter is caused by two components : an interface resistance caused by the presence of a thin interfacial oxide layer between the poly and mono-silicon and a true contact

resistance at the metal-poly interface. Both components are considerably higher than in the case of mono-crystalline silicon emitters. The first one does not exist in a monocrystalline emitter and the second one strongly depends on the surface free carrier concentration. This latter is considerably smaller in the case of a poly-emitter due to carrier trapping and impurity atom segregation at the grain boundaries.

Besides series resistance problems, the thickness of the interfacial oxide layer is not easily controlable which causes serious yield and reproducibility problems in poly-emitter structures. The elimination of this thin layer is therefore two fold advantageous at the expense of reduced current gain.

The reduction of emitter series resistance, coupled with an expected increase in reproducibility is the prime reason for the interest in epitaxial emitter structures. In addition, due to the fact that epitaxial emitters are deposited doped with the appropriate impurity and not doped by compensation, a better emitter efficiency is expected.

2.1. Epitaxial Regrowth of Polysilicon

It has been demonstrated that it is possible to align polysilicon epitaxially to the underlying single crystalline substrate by high temperature furnace annealing or rapid thermal annealing [4-7]. In order to start the alignment process, a direct contact between the polysilicon layer and the single crystalline substrate must, however, take place. In other words the native interfacial oxide, though very thin, should be removed. Figure 1 displays a cross sectional TEM photograph of a partially aligned polysilicon film after a 30' anneal in Argon at 1000°C, and shows clearly how discontinuous the interfacial native oxide layer becomes. This readily occurs at annealing temperatures above 950°C, and can be stimulated by annealing in an oxidizing ambient if the anneal is to be performed at lower temperatures (~ 850°C) [4]. Proper adjustment of the annealing time, temperature and ambient can lead to an epitaxial regrowth of the entire polysilicon film. Undoped, n-type and p-type polysilicon films have shown similar behavior [4-6] which makes low resistivity epitaxial burried p^+ base contacts for self aligned "poly --> epi" bipolar transistors possible.

It is not possible, however, to obtain emitter box-shaped profiles using this method because of impurity redistribution in the aligned region as well as impurity out-diffusion into the substrate during annealing.

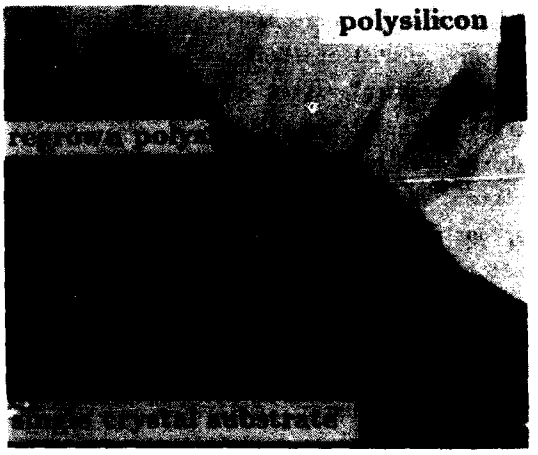


FIGURE 1

Cross sectional TEM photograph showing epitaxial alignment of polysilicon to the underlying substrate after anneal in Argon for 30' at 1000°C (1cm = 80 nm).

2.2. Low Temperature Epitaxial Growth by Glow Discharge Deposition

Another very attractive approach to form a low temperature epitaxial emitter has recently been proposed by the authors [8]. Under proper pre-deposition cleaning conditions, heavily doped amorphous silicon films deposited from a silane plasma at 250°C on a single crystalline silicon substrate will

epitaxially recrystallize throughout the film after an annealing at a temperature of 600-700°C during typically 30'. Figure 2 shows a cross-sectional TEM picture of the as-deposited amorphous silicon film without subsequent heat treatments. This picture clearly shows that over large fractions of the interface the silicon is initially deposited under single crystalline form epitaxially aligned with the substrate. The thickness of this layer depends from point to point and is typically 10 nm. The remaining part of the layer is deposited in the amorphous state. The existance of this thin epitaxially aligned layer is probably due to a combined effect of the presence of F^- ions at the surface resulting from the pre-deposition cleaning step, and the reducing effect of the silane plasma on the "native" oxide during growth. Figure 3 shows the same example after annealing at 600-700°C. Clearly the entire n-type layer has recrystallized up to the surface; the presence of twin defects can be observed. The advantages of this process are the large throughput, the good homogeneity over large areas, the effectiveness of dopant introduction and the low processing temperature. Transistor operation with very reasonable efficiency (emitter Gummel number = 10^{14} $cm^{-4}s$) has been obtained (Fig. 4).

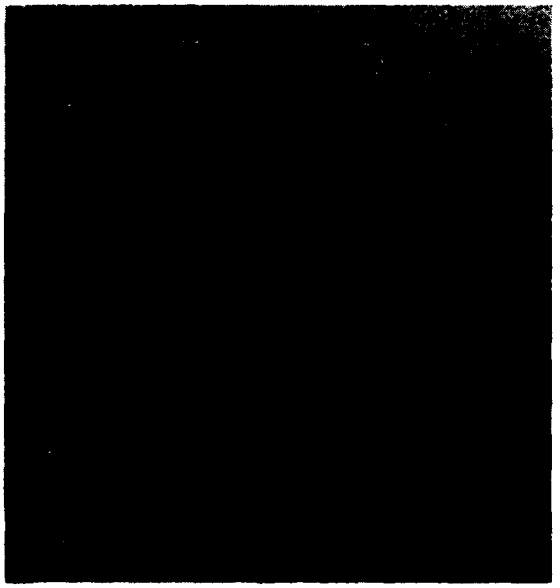


FIGURE 2

HRTEM photograph showing the as-deposited n^+ a-Si:H/c-Si interface. The deposited Si is epitaxially aligned to the substrate over large fractions of the interface.

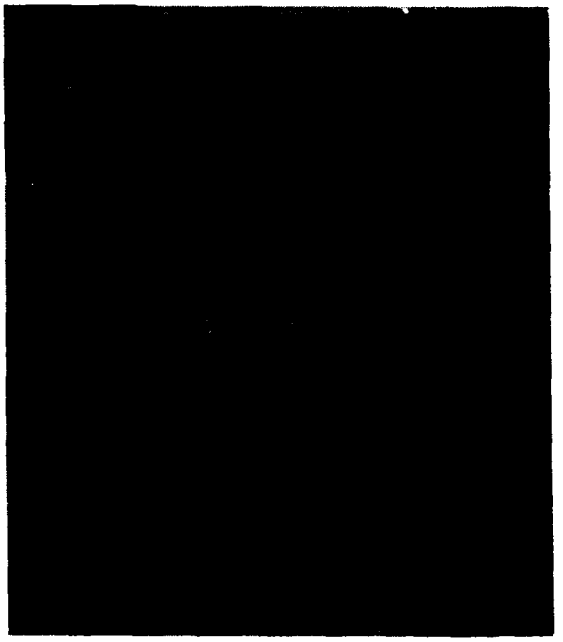


FIGURE 3

HRTEM photograph showing the n^+ a-Si:H/c-Si interface after 30' anneal at 700°C. The entire n^+ a-Si:H film is epitaxially aligned to the substrate. The insert represents the microdiffraction pattern of the recrystallized region.

3. WIDE BANDGAP EMITTER SILICON BIPOLAR TRANSISTORS

Wide bandgap silicon bipolar transistors are heterotype silicon bipolar devices which may allow ultra high f_T ($\geq$ 40 GHz) and room temperature gate delays of 5 ps [9]. These expectations are based upon the following advantages : the combination of a sufficiently large current gain with a low intrinsic base

resistance, the elimination of emitter stored charge and a better control on the uniformity of ultra-shallow box-type emitter base profiles. As discussed earlier these advantages should not be realized at the expense of a large emitter resistance, causing a decrease in transconductance and completely annihilating the expected gain in transistor performance.

At the moment most of the work on silicon hetero-type devices is still in the early development stage since only dc results are largely available; results on the dynamic performance of silicon heterojunction devices arevery limited.

The challenges in wide-gap emitter research can be summarized as follows :

- to find a wide-gap material compatible with state-of-the art silicon processing
- the doping efficiency of the wide gap material must be sufficiently large such that low bulk resistivities and contact resistances can be obtained.

Unfortunately it turns out that these two requirements are not easily compatible. The following materials have been investigated, with variable success, as wide gap emitter :

- GaP : this material is lattice matched to Si but suffers from interface doping effects; so far poor transistor performance has been reported [10].
- SIPOS : yields excellent GN_E but large R_E. Experiments have shown that current gain is not related to the wide emitter gap but to the presence of a thin interfacial oxide layer [11].
- β-SiC : although not lattice-matched on silicon good quality epitaxial growth without large built-in stress has been demonstrated [12]. This material is very promising, although device performance has not yet been reported.
- Amorphous Si (a-Si:H) : excellent GN_E but large R_E (> 10^{-3} Ωcm^2) [13]; large series resistance is caused by contact rather than by the presence of an interfacial oxide layer. The TEM picture of Fig. 2 demonstrates the onset of epitaxial alignment over a large fraction of the interface at the growth temperature of 250°C in the plasma CVD reactor. This indicates that the large emitter Gummel number is not an interface but a true heterojunction effect. Amorphous SiC:H has also been tried [14], but devices suffer from similar problems as with a-Si:H.
- Microcrystalline silicon [15] : allows large bandgap (1.4 eV) and reasonably low resistivities (10^{-2} Ωcm) to obtain low series resistance. However, growing this material reproducibly using the conventional plasma CVD without epitaxial alignment may be difficult.

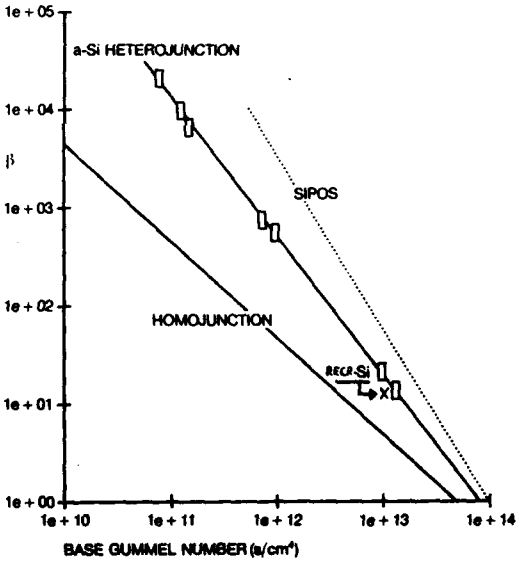


FIGURE 4

Max DC current gain versus base Gummel number for different types of emitters.

Figure 4 presents some experimental results indicating an increase in the dc current gain β when Si wide bandgap emitters are used. From these preliminary results it follows that, in the opinion of the authors, two

"wide-gap" materials should be considered as serious candidates : SiC and to a smaller extent microcrystalline silicon.

4. NARROW BASE SILICON BIPOLAR TRANSISTORS

It is now well known [16] that the Ge_xSi_{1-x} layers can be grown pseudomorphically on silicon if the layers are thinner than the critical thickness. If such a p-type layer is overgrown by an n-type doped Si layer a narrow base transistor is formed. Initial calculations indicate that with x = 0.2 a large bandgap difference between emitter and base and a critical thickness large enough to avoid punchthrough and yield low base resistance, can be combined. As in this structure the silicon in the emitter will remain cubic and unstressed the bandgap difference will be almost completely seen as a bandoffset at the valence band edge yielding a very efficient np heterojunction.

In the opinion of the authors such a *narrow-gap* base heterotransistor is extremely promising for the following reasons :

- emitter and collector interchangeability
- smaller turn-on voltage, allowing a somewhat lower power dissipation
- due to the absence of collector stored charge a high speed saturated logic can be developed
- no problems with low emitter specific contact resistance.

5. CONCLUSIONS

Directly diffused or implanted emitters are not compatible with very thin base VLSI bipolar transistors. Polysilicon emitters suffer from high emitter resistance and in the short term epitaxial emitters seem to be the best alternative. On the other hand, more research sould be oriented towards narrow bandgap base heterojunction bipolar transistors as they present comparably high emitter efficiency to wide band gap emitters with, however, no emitter resistance problems.

6. ACKNOWLEDGEMENT

The authors would like to thank J. Vanhellemont for the TEM analysis.

REFERENCES

[1] Crabbé, E., Swirhun, S., del Alamo, J., Pease, R.F. and Swanson, R.M., IEDM Tech. Dig. (1986) 28.

[2] Stork, J.M.C. and Cressler, J.D., Symp. on VLSI Technol. Dig. Tech. Papers (1986) 47.

[3] Solomon, P.M., Proceedings of the IEEE 70 (1982) 489.

[4] Ghannam, M.Y. and Dutton, R.W., Proceedings of the IEEE Bipolar Circuits and Technology meeting (1986) 5.

[5] Tsaur, B.Y. and Hung, L.S., Appl. Phys. Lett. 37 (1980) 648.

[6] Wong, C.Y., Michel, A.E., Isaac, R.D., Kastl, R.M. and Mader, S.R., J. Appl. Phys. 55 (1984) 1131.

[7] Natsuaki, N., Tamura, M., Miyazaki, T. and Yanagi, Y., IEDM Tech. Dig. (1983) 662.

[8] Baert, K., Symons, J., Vandervorst, W., Vanhellemont, J., Caymax, M., Poortmans, J., Nijs, J. and Mertens, R., to be submitted to APL for publication.

[9] Wieder, A.W., IEDM Tech. Dig. (1986) 8.

[10] Kroemer, H., Proceedings of the IEEE 70 (1982) 13.

[11] Oh-uchi, N., Hayashi, H., Yamoto, H. and Matsushita, T., IEDM Tech. Dig. (1979) 522.

[12] Yamanaka, M., Daimon, H., Sakuma, E., Misava, S. and Yoshida, S., J. Appl. Phys. 61 (1987) 599.

[13] Ghannam, M., Nijs, J., Mertens, R. and De Keersmaecker, R., IEDM Tech. Dig. (1984) 746.

[14] Sasaki, K., Furukawa, S. and Rahman, M., IEDM Tech. Dig. (1985) 294.

[15] Ghannam, M., Nijs, J., De Keersmaecker, R. and Mertens, R., EPS abstracts, ESSDERC (1985) 46.

[16] Jorke, H. and Herzog, H.-J., J. Electrochem. Soc. 133 (1986) 998.

Session B2.2

SOI Workshop II

Chairman: D. Mc Caughan

Tuesday, September 15, 1987

ELECTRICAL PARAMETERS OF SOI MATERIAL OBTAINED BY ZMR AND OXIDIZED POROUS SILICON

M. Haond, G. Bomchil, J-L. Regolini, D. Bensahel, D. Dutartre, D-P. Vu, K. Barla, H. Halimaoui, R. Herino.

Centre National d'Etudes des Telecommunications, BP 98, 38243 Meylan Cedex, France

A. Monroy, S. Thouret, Y. Gris.

Thomson Semiconducteurs, BP 217, 38019 Grenoble Cedex, France

Lamp-Zone Melting Recrystallization (ZMR) of deposited silicon on oxide has proved to be suitable for making devices. We present here electrical results obtained in this material on batches of 4-in. wafers, which confirm its crystalline quality. We also present recent results obtained in SOI material prepared by oxidizing a buried porous layer. Since laser-ZMR is still in the race toward the fabrication of a material compatible with 3-D circuits fabrication, some new results are periodically available. By presenting electrical results of the three types of material, we compare and discuss the future trends in SOI concerning each of these three techniques.

1. INTRODUCTION

Silicon-On-Insulator (SOI) is expected to help solving some of the present problems encountered in the CMOS technology of VLSI circuits, namely in the race toward smaller dimensions: latch-up, dielectric isolation and density of integration, power consumption,... Many techniques have been investigated. The use of an oxygen ion implantation (SIMOX) is presently widely studied over the laboratories, and is discussed elsewhere in these proceedings. We have been working on Zone Melting Recrystallization (ZMR) for a long time /1/. This technique has therefore reached a level where batches of 4-in. wafers can be processed reproducibly and provide a device-worthy material /2/. Another technique is attractive since the processing apparatus is simple and cheap: the oxidation of porous silicon obtained by anodizing a N/N+/N structure. The last technique we report below is laser-ZMR which is one of the techniques which should be used for making 3-D devices. We present these three techniques and compare their respective electrical parameters as measured after having ran wafers of each type in a 3µm CMOS technological process.

2. SOI MATERIAL PREPARATION

The preparation of lamp-ZMR SOI material has already been reported /3/. Shortly, a grating is etched in an oxidized wafer. It consists in 0.4 µm deep and 36 µm wide stripes separated by 4 µm lines. A 0.5 µm to 0.6 µm thick encapsulated polysilicon film is used. The cap is a 1.6 µm thick oxide. It avoids the delamination of the liquid silicon and limits mass transport at melting. The lamp apparatus /3/ consists in a row of halogen lamps used to preheat the wafer up to 1150 °C. An additional lamp placed in an elliptical mirror is used to melt the deposited polysilicon film. The molten line is scanned at a speed of 0.2 mm/s. The grating etched in the underlying oxide provides an effective defect entrainment, thanks to a solidification front modulation /4/. Fig.1a presents a SEM micrograph of a cross-section of a typical lamp-ZMR wafer. Extended characterizations of these samples have already been reported /5/. The remaining defects such as grain-,subgrain boundaries (GBs, SGBs) or precipitates are entrained upon the 4 µm steps of the relief grating. The wafers are flat and the slip lines if there are any are located in the bulk substrate and not in the SOI film itself. Batches have been processed which

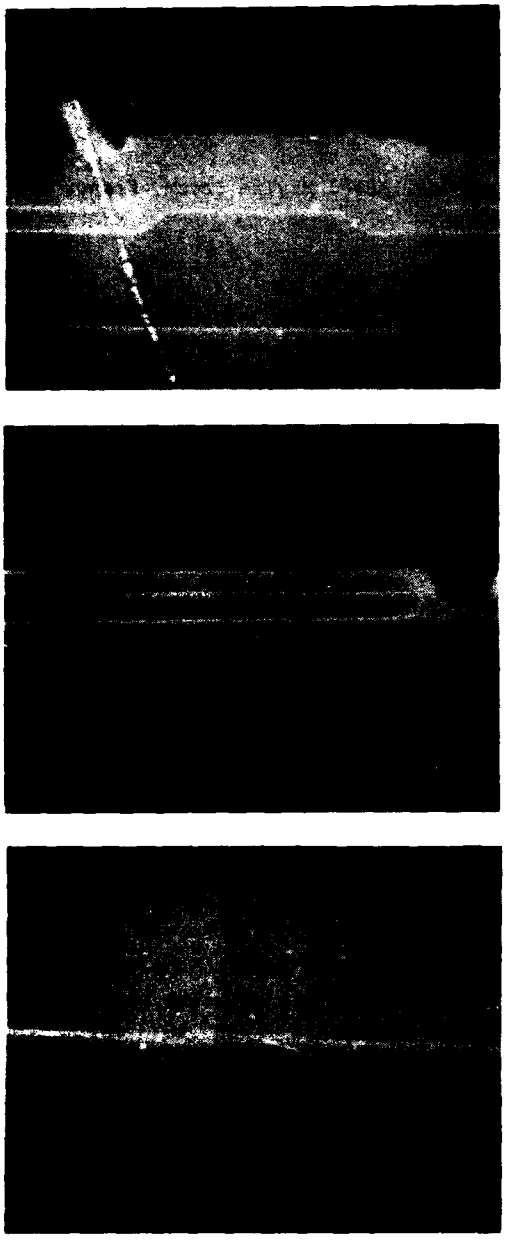


FIGURE 1

SEM cross-sectional micrographs of:
a) a lamp-ZMR oxide capped sample near a defect entrainment line
b) an oxidized porous sample near an anodization access window . The sample has been etched in a 1:7 HF:FNH4 solution
c) a laser-ZMR sample near a SEG-filled seed. The sample has been immersed in a buffered HF solution.

show the reproducibility of the lamp-ZMR procedure as well as the compatibility of the wafers with a CMOS process line.

A different approach has been studied to fabricate SOI material: it is based on the transformation of single crystal silicon into porous silicon by anodization in an HF electrolyte. By taking advantage of the preferential anodization of N+ silicon, a buried porous layer can be obtained by opening windows in the N-type epitaxial layer of an N/N+/N structure /6/. The formation of porous silicon proceeds laterally until the whole volume of the buried layer is transformed, in a self-limited reaction. We have been working with a controlled potential in a potentiostatic three electrodes configuration. The anodization conditions in HF electrolytes solutions are adjusted to give an homogeneous porous layer of about 56% porosity. An oxidation step provides a buried oxide which is equivalent to a thermal oxide /7/, regarding its resistance to chemical etching in HF: FNH4 solutions. Fig.1b is a SEM micrograph of a cross-section of the SOI strucure after the porous layer has been oxidized. The buried oxide is homogeneous, displays flat and abrupt interfaces. In this sample, the N+ layer doping level was 1.5 E19 cm-3 (antimony) and was anodized in 35% HF solutions in ethanol. The oxidation procedure has been described elsewhere /7/. By using this procedure, 4-in. SOI wafers have been fabricated which are compatible with a standard CMOS technological process. The SOI material consists in 4/36 µm window/active region stripes, all across the wafer.

Laser-ZMR has been a pionneering technique for the obtention of SOI material, but its application has been restricted because of several drawbacks. The spot size and related overlapping problems remain. Since the two previous techniques cannot afford the opportunity of making 3-D devices, the laser has not been abandonned. Therefore, new efforts have been made on the way toward obtaining large defect-free areas. In

laser-ZMR where a seed is used, defects are related to the thermodynamical behaviour of the seed area upon melting, i.e. the volumetric contraction of silicon and the temperature gradient between the seed and the SOI region. These effects can be overcome by using a seed filled with single crystal Si or by using discontinuous seeds /8/.

In our case, we have used a Selective Epitaxial Growth process (SEG) /9/. Our samples are as follows: 4-in. wafers are oxidized up to 0.5-0.6 µm and then etched by Reactive Ion Etching (RIE) for delineating the seed regions: 2 or 4 µm wide and 40 µm pitch. The seed lines are <100> or <110> oriented. The solidification front is here controlled by the trailing edge of the molten spot.

A solidification front parallel to the <110> direction is therefore obtained by using an elliptical spot slanted at 30 ° from the <100> scan direction.

A 1 µm poly-Si film is then deposited, followed by the deposition of a 1 µm oxide cap. Since the slanted elliptical spot is scanned parallel to the seed lines, the defects are rejected near one end of the seed. Fig.1c shows a SEM micrograph of a cross-section of our sample near the seed region. The high crystal quality has been confirmed by TEM. The crystalline defects are the SGBs which correspond to about 1 degree of misorientation between each side of the SGB. Similar results have been obtained using discontinuous seed /8/. Isolated dislocations exist in the seed regions. They do not extend very deep into the substrate and their density is lower with 2 µm seeds than with 4 µm openings.

3. ELECTRICAL PARAMETERS

The three types of SOI wafers have been processed in similar self-aligned 3 µm poly-Si gate CMOS sequences. The main technological parameters are summarized in Table 1. The laser results presented below have been obtained on samples prepared and recrystallized by the LETI and using discontinuous seeds /8/. Natural transistors are

SOI	Gate Oxide	Lateral Isolation	Gate Material
5000 A	420 A	RIE of mesas	Poly-Si:4200 A

Table 1 : Process Parameters.

fabricated by avoiding the channel implants, i.e. using the as-prepared SOI material. They can either be enhancement or depletion mode transistors. These natural transistors provide an adequate tool to study the properties of the interfaces and to measure the residual doping concentration, by C-V, I-V and transconductance (gm) measurements /10/. Since the SOI film thickness is about 5000 A, the channel will be completely depleted or not, depending on the residual doping level of the SOI. Table 2 presents values of drain current as measured on depletion N+/N/N+, W/L= 30/30µm (natural) transistors at Vg= -2V and Vd= 5V. It is below E-13 A/µm in the lamp-samples whereas it is about E-9 A/µm in the laser-samples and E-6 A/µm in the FIPOS samples. These values are attributed to the residual doping level and to interface states at the back interface. In the three types of SOI material, the residual doping is N-type. It is in the low E15 cm-3 in the lamp samples and in the high E15 cm-3 in the laser samples. No satisfactory explanation has been found to explain this N-type, whereas it is in the E16 cm-3 for the FIPOS samples where it can be attributed to a diffusion or to any incorporation mechanism of species from the N+ layer used for the preparation of the material (N/N+/N structure). In order to better qualify the material quality, we have derived the minority carrier lifetime from the time to form the inversion layer in a depletion mode transistor /11/. Table 2 summarizes the orders of magnitude obtained in each type of SOI. It shows that the lamp material is approaching bulk material quality whereas FIPOS and laser materials have traps which reduce the lifetime. As shown in Table 2, we have also derived values of electron and hole mobilities from gm(Vg) curves at Vg=

	Residual Doping	tg	μe	μp
Lamps	N : 2-3 E15 cm-3	10 μs	900	230
FIPOS	N : 2-3 E16 cm-3	1 μs	500	145
Laser	N : 7-8 E15 cm-3	1 μs	1000	200

tg: Minority Carrier Lifetime.
μe, μp: Electron, Hole Mobility.

Table 2 : Electrical Parameters Obtained from Natural Transistors (N+/N/N+ and P+/N/P+).

Vfb (flatband voltage) /12/. Notice that these values do not correspond to those measured in inversion or accumulation layers. They are close to the bulk values /13/ whereas the second ones are lower as a result of increased scattering.

These results are confirmed by measurements performed on enhancement mode transistors (where the channel has been ion implanted). In Table 3, we present leakage current levels as measured on 72/5 μm N-type edgeless transistors at Vg= -2V and Vd= 5V. In each type of SOI, they remain below 1 pA/μm of channel width. Table 3 also displays surface mobility values as derived from Id-Vg curves in the linear region. The high value of the electron mobility together with a low hole mobility in the laser-samples could be due to some residual local stress within the SOI film. The low hole mobility of the FIPOS sample could result from a poorer quality either of the interfaces or of the material as compared to the other materials. These parameters have been derived from measurements performed on edgeless transistors since these allow to avoid the parasitic edge channel encountered in mesa etched N-type transistors /14/. Fig.2 shows Id-Vg curves of both P- and N-type edgeless 72/5 μm transistors obtained in lamp-ZMR wafers at Vd= 0.1 V. The subthreshold slope is about 120 mV/decade and confirms a low density of interface states. This slope can overcome the theoretical "ideal" value as a result of the "kink" effect /15, 2/due to the floating substrate /16/. It can be avoided by connecting the source and the channel or reduced by thinning the SOI film /17/.

	Leakage Current (•)	μe cm2/V.s	μp cm2/V.s	tp ns
Lamps	<pA	620	220	1
FIPOS	<pA	570	155	1.7
Laser	<pA	670	185	2.2

(•) Measured at Vg=+/-2V ; Vd=+/-5V.
μe, μp: Electron,Hole Mobilities.
tp: Propagation Delay Time per Gate (3.5 μm effective channel length).

Table 3 : Electrical Parameters Obtained from N- and P-Enhancement Mode Transistors .

249-stage, 3.5 μm effective gate length ring oscillators have been fabricated, in order to test the dynamic characteristics of the materials. Table 3 summarizes the results for the three SOI techniques. We believe that in the laser-samples where the mobility on individual transistors is high (see Table 3), the increased propagation delay time is partly due to the thin oxide used to isolate the seeds (gate oxide), and also to the low thickness of the underlying oxide (0.3 μm).

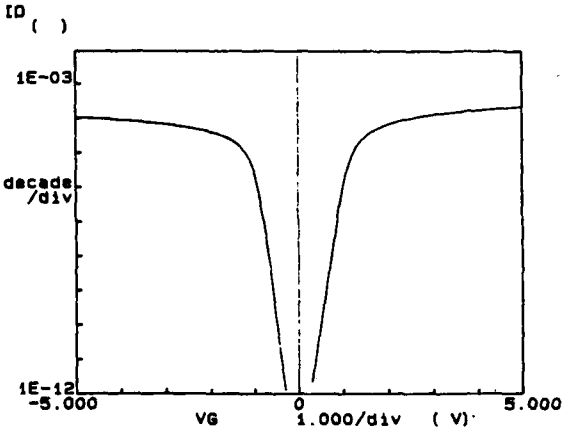


FIGURE 2

Id-Vg curves of P- and N-channel transistors fabricated in lamp-ZMR material. Channel width, W= 72um and gate length, L= 4um. Vd= +/- 0.1V and Vg= -/- 5V. The Bulk substrate voltage Vb is 0V.

The number of processed wafers is too low for the laser- and FIPOS-SOI material to provide reliable statistical studies from a wafer to another or from a run to another. It must be pointed out, however, that in single wafers the laser- and FIPOS-SOI film thickness is very uniform. This point is very important for technological steps such as SOI film thinning, dielectric isolation (mesa etching or LOCOS definition), energy adjustment of the Drain/Source implants...

Statistical results are available for the lamp-SOI material where many batches have been processed. They have first shown that the residual doping level should be reduced down to the E15 cm-3 or less in order to allow the adjustment of the threshold voltage of both N- and P-type transistors. Notice in Fig.2 the symmetrical characteristics of both P- and N-channel transistors. In Fig.3, we display the threshold voltage distribution of enhancement mode 20/3 µm N-type transistors obtained on 4 4-in. wafers of a same batch. The mean value is 0.9V with a standard deviation of 61 mV. We have already presented some statistical results, namely on the influence of the defect entrainment lines and shown how they scatter the threshold voltage distribution /2/.

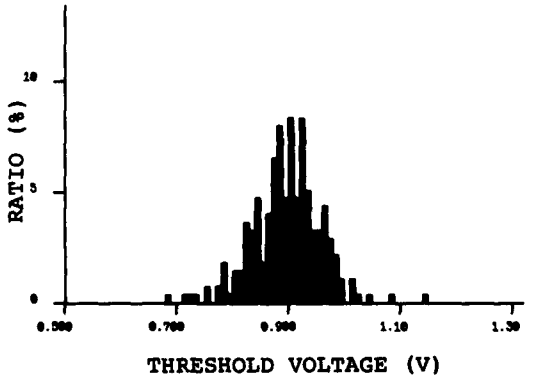


FIGURE 3

Threshold voltage distribution of N-type transistors fabricated in lamp-ZMR wafers. W= 20um, Leffective= 1.7um. The mean value is 0.9V and the standard deviation 61 mV.

4. DISCUSSION

We have presented above the main electrical results obtained in each of the three techniques. It must be pointed out that these results are not independant of the level of maturity of the technological device process for a given material. Since lamp-material has reached a process-compatible quality before the other two, they are a step forward on their way toward the production of SOI circuits for the microelectronics. Specific technological steps necessary. Moreover, the design constraints brought on by the relief grating have been taken into account by the circuit designers who use the defective silicon for highly doped parts of circuits. Mass-transport still has to be improved if one wants to thin the SOI film down to hundreds of angstroms. This should be useful for avoiding the "kink" effect without density loss. Another "effect" has to be reduced, that is the previously reported "shrink" of the lamp-ZMR wafers /3/. It leads to a narrowing of the defect lines. It must be quantified precisely and be reproducible in order to have it correctly corrected at the first mask level, that is the relief grating level.

On the other hand, laser- and FIPOS-SOI materials are quite new, as far as the VLSI technological process is concerned. FIPOS is very attractive since the useful material is similar to the starting bulk crystal as it does not undergo any transformation. However, the residual doping level should be reduced at least of one order of magnitude if we want to adjust the threshold voltages. Moreover, new design constraints are brought in by the fact that there is no silicon left in the access windows of the N/N+/N structure after oxidation of the porous buried layer. The interfacial quality of the oxide should also be further improved. Finally, the laser technique has to be characterized more specifically and precisely, namely concerning the residual crystalline defects (SGBs, twins, stacking faults,...) and the residual stress within the recrystallized material. Moreover, the long term stability of the laser also has to be certified. We believe, however, that this technique should find some specific applications in the field of 3-D circuits.

To summarize, we believe that it might be possible to classify the three techniques in a chronological order of applications to circuit fabrication. Lamp-ZMR material is ready for circuit development. FIPOS-material should find a medium term application, after the main residual problems indicated above will be solved. Laser-ZMR has a certified but probably long term 3-D circuit field of application and should still demonstrate its reproducibility and

crystalline quality on a large scale if it wants to be applied shortly in the VLSI 2-D technology.

5. CONCLUSIONS

We have shown that among three different techniques used in the recent years to fabricate SOI material, i.e. lamp- or laser-ZMR and FIPOS, the level of advancement is quite different in both material and technological processes. Statistical and electrical parameters obtained in lamp-ZMR batches of wafers show that this technique is ready for providing substrates for making CMOS devices. On the other hand, problems related to remaining defects in the material or to a high residual doping level still have to be solved in the other two techniques before they can provide reliable substrates.

ACKNOWLEDGMENTS

We wish to thank G. Gimine and the Pilot Line at Cnet who took an important part in the technological process. We acknowledge M. Montier and the CMS department from Thomson Semiconducteurs for his encouragement in this work.

This work was partly supported by an Esprit Project.

REFERENCES

/1/ Haond, M. and Vu,D-P., Electron. Lett. 18 (1982) 727.

/2/ Haond, M., Vu, D-P., Monroy Aguirre, A., Perret, S., Circ. and Dev. IEEE,(1987) in Print.

/3/ Haond, M., Dutartre, D. and Bensahel, D., MRS-Europe Proceedings (1985) 417.

/4/ Dutartre, D., Appl. Phys. Lett. (1986) 350.

/5/ Haond, M., Dutartre, D., Bensahel, D., MRS Symp. Proc. Vol.53 (1986) 83.

/6/ Holmstrom, R.P., Chi, J.Y., Appl. Phys. Lett., 42 (1983) 386.

/7/ Barla, K., Bomchil, G., Herino, R., Monroy, A., Gris, Y., Electron. Lett. 22 (1986) 1291.

/8/ Mermet, J.L., Achard, H., Bono, H., Joly, J.P., to be published in ESPRIT Technical Week, Bruxelles, Oct. 87.

/9/ Regolini, J.L., Dutartre, D., Bensahel, D., Karapiperis, L., Garry, G., Dieumegard, D., Electron. Lett. 23 (1987) 495.

/10/ Vu, D-P. and Pfister, J.C., Appl. Phys. Lett. 48 (1986) 50.

/11/ Vu, D.P. and Pfister, J.C., Appl. Phys. Lett. 47 (1985) 950.

/12/ Vu, D-P., Chantre, A., Mingam, H. and Vincent, G., J. Appl. Phys. 46 (1984) 1682.

/13/ Vu, D-P., Chantre, A., Ronzani, D. and Pfister, J.C., MRS Symp. Proc. Vol.53 (1986) 357.

/14/ McGreivy, D., Electron. Dev. ED-24 (1977) 730.

/15/ Davis, J.R., Glaccum, A.E., Reeson, K., Hemment, P.L.F., Electron Dev.Lett. EDL-7 (1986) 570.

/16/ Tihanyi, J. and Schlotterer, H., Solid State Electron. 18 (1975) 309.

/17/ Colinge, J.P. and Kamins,T.I., in print.

POROUS ANODISED SILICON FOR FULL DIELECTRIC ISOLATION: The Development of an n/n+/n Device Route

D Brumhead, J G Castledine and J M Keen
J M Cole, L G Earwaker, J P G Farr, P E Grzeszczyk, J L'Ecuyer, M Loretto and I M Sturland*

Royal Signals and Radar Establishment, St Andrews Road, Great Malvern, Worcs WR14 3PS, UK
* University of Birmingham, P O Box 363, Birmingham B15 2TT, UK

The n/n+/n route to porous silicon has been used to produce fully dielectrically isolated silicon islands. Results are presented to show that doping of the island with residual n+ material can be avoided and that the silicon is of high crystalline perfection. The technique is shown to avoid the limitations of the original p-n technique and to be extremely promising for SOI Device applications.

1. INTRODUCTION

Progress is reported in applying oxidised porous anodised silicon for full dielectric isolation in VLSI. Early procedures [1] were based on selectively anodising p-type regions, and NTT made a 64K RAM using this technique. The method, however, had limitations with respect to island width and crystallographic perfection, and resulted in thick porous layers. Inevitably, a spike was left under the silicon islands. Other approaches involve growing good quality molecular beam epitaxial layers on porous silicon [2] or rely on a buried layer beneath the silicon wafer surface to direct the current flow laterally, under the device islands [3]. Both of these methods have disadvantages in practice.

We have explored n/n+/n enhanced lateral anodisation as an alternative route [4]. Its basis is that the suceptibility to anodising is dependent on the n-type doping level, so that higher dopant concentrations lead to lower anodising voltages and lower porous densities. Therefore n+ material can be anodised preferentially leaving a lower-doped n-type island. The advantages of this technique over using p/n selectivity are that the thickness of the porous silicon and hence the buried oxide can be decoupled from the width of the islands, there is no residual spike beneath the centre of the island, therefore the islands as formed and after oxidation are potentially defect free. A disadvantage is the need for a thin epitaxial layer with an abrupt interface to a buried n+ layer.

2. EXPERIMENTAL

A strip mask has been used to produce wafers with different structures in each of four quadrants. Island strips of width 20, 30, 40 and 60 microns have been produced with 5 micron n+ anodising entry windows between them in the four quandrants, respectively. The wafers were anodised at constant voltage in an ethanoic/HF electrolyte to convert the n+ silicon both between and under the islands into porous silicon about half a micron thick. The wafers were subsequently oxidised and analysed.

3. RESULTS

Transmission electron micrographs (TEM) of cross-sections of island-doped wafers (Figure 1) illustrate the isolation obtained, the retention of the island geometry and the good crystallographic quality of the island.

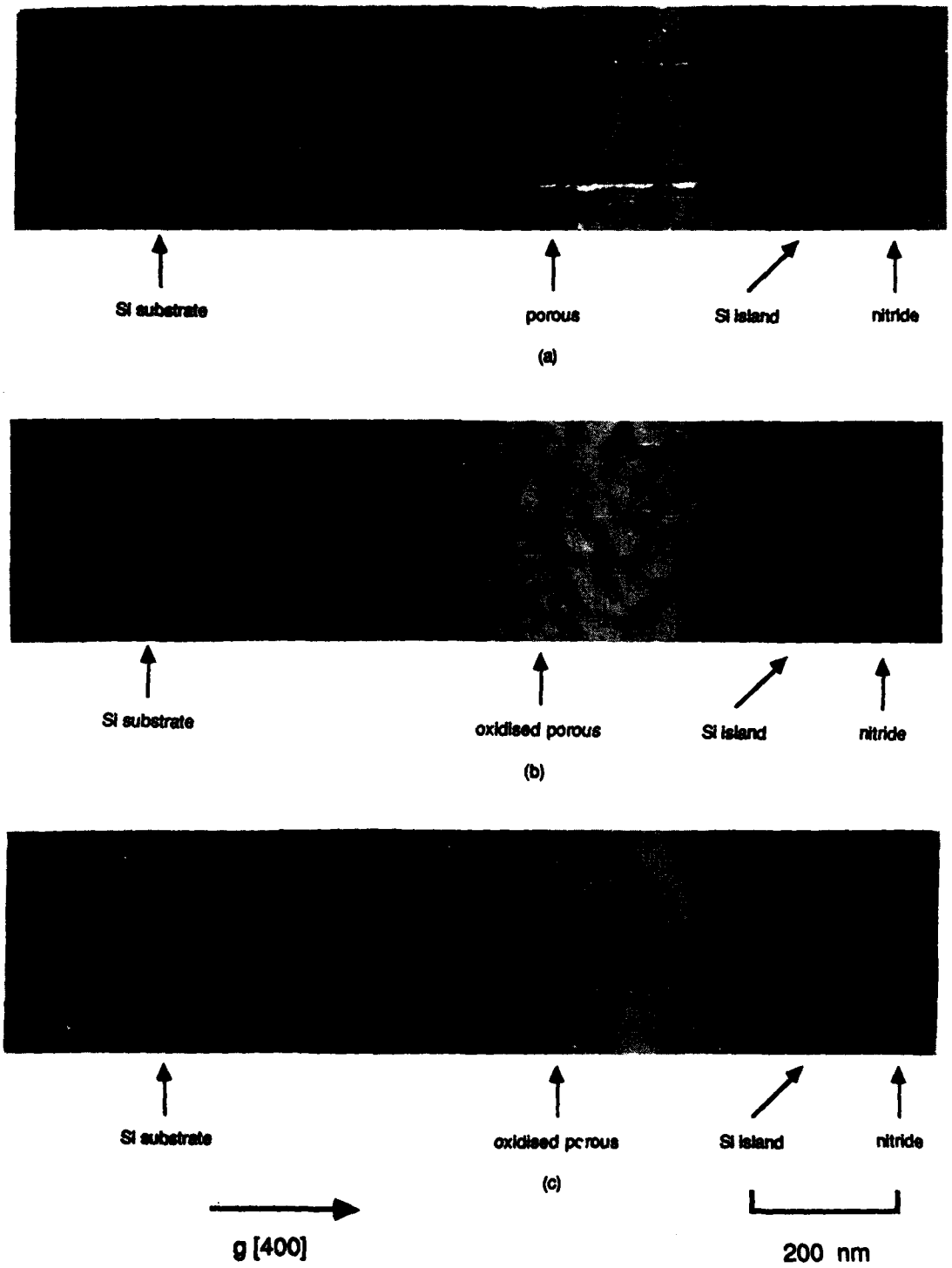


Figure 1. Cross-section TEM micrographs showing the n/n+/n microstructure following (a) anodising and (b) oxidation at 300°C for 1 hour and 800°C for 2 hours, (c) 300°C for 1 hour, 800°C for 1 hour and 1090°C for 4 minutes followed by an anneal in dry nitrogen for 1 hour. Note the absence of defects in the Si island and the sharpness of the interface between the Si island and the porous Si.

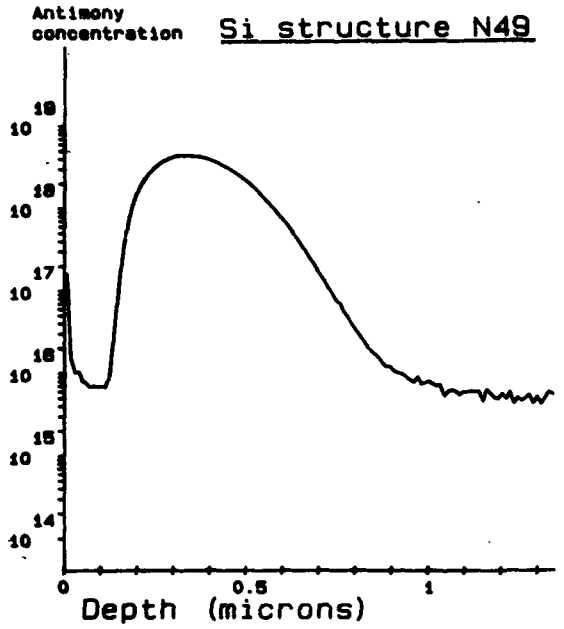


Figure 2(a). SIMS profile showing the initial dopant distribution through a vertical section of the island structure.

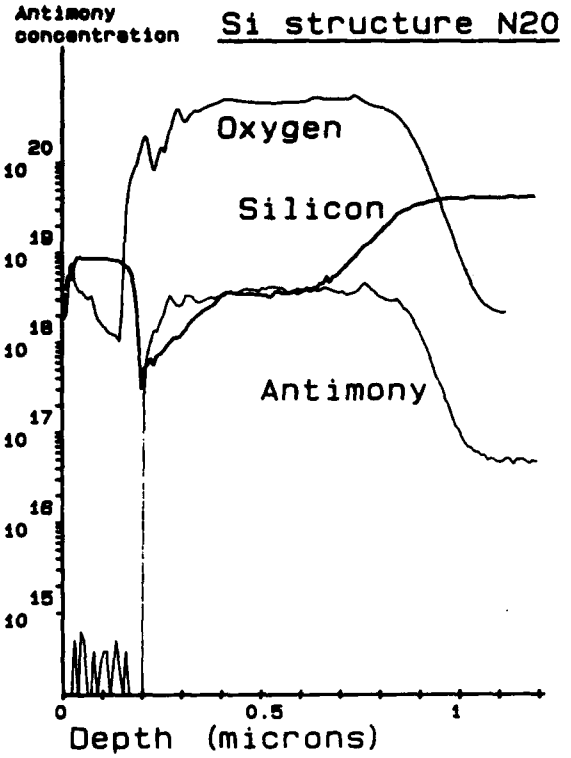


Figure 2(b). SIMS profile showing dopant distribution after the oxidation of the porous silicon.

The non-uniform structure of the as-anodised porous silicon is related to the implant profile which is well-revealed by SIMS analyses (Figure 2a). The pore size and structure is determined by a complex relationship between dopant concentration, electrolyte composition and the anodising current density. By careful choice of conditions, the doping levels and current density effects can be used to counteract each other in order to achieve a much better uniformity than that shown.

The steepness of the profile determines the roughness of the interfaces between the silicon island/porous silicon, and substrate/porous silicon. On oxidation, the porous silicon, together with some of the silicon at the back of the island and at the substrate interface, is converted to SiO_2. This is shown in Figure 1 (TEM) and Figure 3 (Rutherford Backscattering).

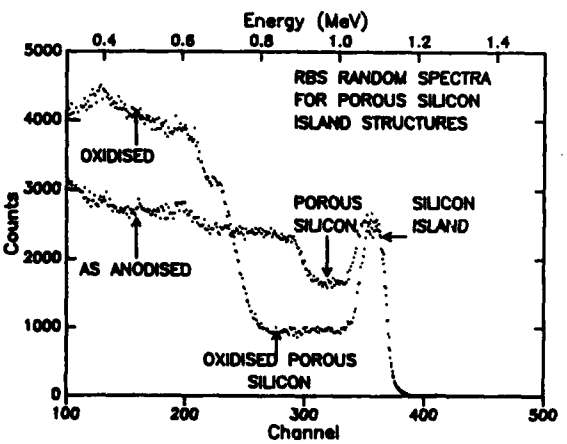


Figure 3. Rutherford backscattering (RBS) spectra from randomly oriented samples of as-anodised and anodised-plus-oxidised (1 hour at 300°C plus 2 hours at 800°C) porous silicon. The spectrum from the oxidised porous sample indicates fully oxidised SiO_2 with some of the back face of the island also being oxidised.

If the anodising has been done correctly, the oxidation of the whole of the back of the island is uniform. Due to the steepness of the impurity profile and the low oxidation temperatures involved, an oxide barrier can be created which makes it possible to avoid doping the back of the island (Figure 2b).

4. CONCLUSIONS

Very lightly doped islands have been produced with sharp interfaces, good island morphology and low defect densities. The n/n+ route has been shown to overcome the limitations associated with the original p/n porous silicon technique. It is therefore, an extremely promising approach to SOI device production.

ACKNOWLEDGEMENTS

This work was partially funded by the UK Alvey programme.

The assistance of Plessey Research (Caswell) Ltd and the GEC Hirst Research Centre in the preparation of these samples and of Mr Blackmore (RSRE) for the SIMS analyses is gratefully acknowledged.

REFERENCES

[1] K Imai, Solid State Electronics, Vol 24 (1981) pp 159-164.

[2] G Konaka, M K Tabe and T Sakai, Appl Phys Lett 41 (1983) 86.

[3] J Benjamin, J M Keen, A G Cullis, B Innes and N G Chew Appl Phys Lett 49 (12) 22 September 1986, 716.

[4] R P Holmstrom and J Y Chi, Appl Phys Lett 42 1983 386.

© Copyright at HMSO, London, 1987

STACMOS: A BASIC 3-DIMENSIONAL CMOS PROCESS

R. Buchner, K. Haberger, P. Seegebrecht and P. Panish

Fraunhofer Institut fur Festkörpertechnologie
Paul-Gerhardt-Allee 42
D-8000 Munich 60, West Germany

MOS Transistors have been fabricated in two independent active device layers, the second of which has been formed through laser recrystallization of a thin polysilicon layer. The effect of the fabrication process on the devices in the silicon substrate has been investigated and characterized through electrical measurements.

1. INTRODUCTION

Integrated circuit semiconductor development is characterized by constantly increasing device packing density. Until recently this has been achieved by reducing the lateral device dimensions while increasing the chip area. The magnitude of reduction in device dimensions has proved to introduce significant technological problems. Another possibility for increasing the level of integration is the utilization of one or more additional device layers. In addition to reducing the interconnection length such a tactic allows the realization of completely new circuit concepts as well as making possible the utilization of mixed technologies.

In order to investigate the effect of recrystallization of the upper layer of a 3-dimensional circuit on the devices in the underlying layer a 3D-CMOS process has been developed.

2. PROCESS

Since the primary purpose of this investigation is to determine the effects of the recrystallization process, the technology has been conceptualized to be as simple as possible. Two active device layers have been fabricated, the first of which is in the monocrystalline silicon substrate, and the second which is fabricated in a thin recrystallized polysilicon layer. This arrangement suited itself to a CMOS circuit structure in which the NMOS and PMOS devices are fabricated in their own layers respectively thus simplifying the process, though it is in principle not necessary. Due to the fact that arsenic is the most thermally stable element in silicon it was decided that the n-channel devices would be fabricated using As doped source-drain regions in the monocrystalline silicon substrate.

The 3D process developed requires 10 mask steps and utilizes virtually exclusively standard semiconductor processes. Starting material for the process is 3" p-type (100) silicon wafers. The first phase of the process is the utilization of a standard poly-gate MOS process to fabricate the n-channel devices using a 500 Å thick gate oxide with a minimum feature size of 4 um. The source-drain regions are As doped using ion implantation. At the completion of this phase the surface exhibits significant topography. It is known that such nonplanar features have undesirable effects on the recrystallization process resulting in nucleation sites for grain boundary formation [1]. These effects are minimized by applying a planarization process which serves simultaneously to insulate the first device layer. A two stage planarization process is used. In the first stage a LPCVD oxide is deposited and then patterned to fill the depressions in the

source-drain areas. A low temperature oxide (LTO) is used for which the etch rate in buffered HF is roughly three times that of a thermal oxide thus allowing processing without special etch stops. With the proper mask design this process allows a nearly planar surface to be obtained. The remaining irregularities are minimized in a reflow planarization step in which a 0.1 um oxide layer is deposited followed by 0.4 um of Phosphorous-Silicate-Glass (PSG). After reflow a 0.2 um oxide layer is deposited to serve as a diffusion barrier.

After completion of the devices in the first layer a 0.5 um thick layer of polysilicon is deposited, doped, and covered with an 850 A thick LPCVD oxide capping layer. The polysilicon is recrystallized over the entire surface with the help of an argon laser. This is necessary in order to avoid excessive absorbtion in the underlying monocrystalline substrate. The laser power, beamwidth, scan speed and substrate temperature are 12 W, 70 um, 3 cm/sec and 500 °C respectively. Modification of the melt-zone temperature profile is achieved through beam forming as shown in Figure 1. Following recrystallization the silicon layer is divided into individual islands to allow for contacts to the devices in the substrate layer.

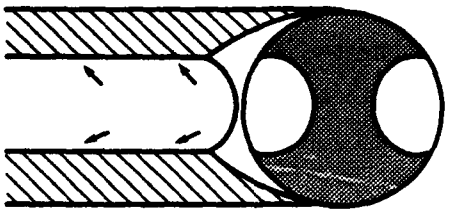


Fig. 1: Beam form of the argon laser.

P-channel polysilicon gate MOS transistors with a 500 A thick gate oxide are formed in the recrystallized silicon layer using a modified standard MOS process. A substrate contact for the devices in the second layer is included since it has been found that the characteristics of devices with a floating substrate are different than those at a fixed potential. Additionally the fixed substrate allows for minimization of coupling effects between the two layers.

The contact window opening proceeds using a combined dry/wet etch process which takes advantage of the high selectivity of wet chemical etching while obtaining low levels of lateral underetching.

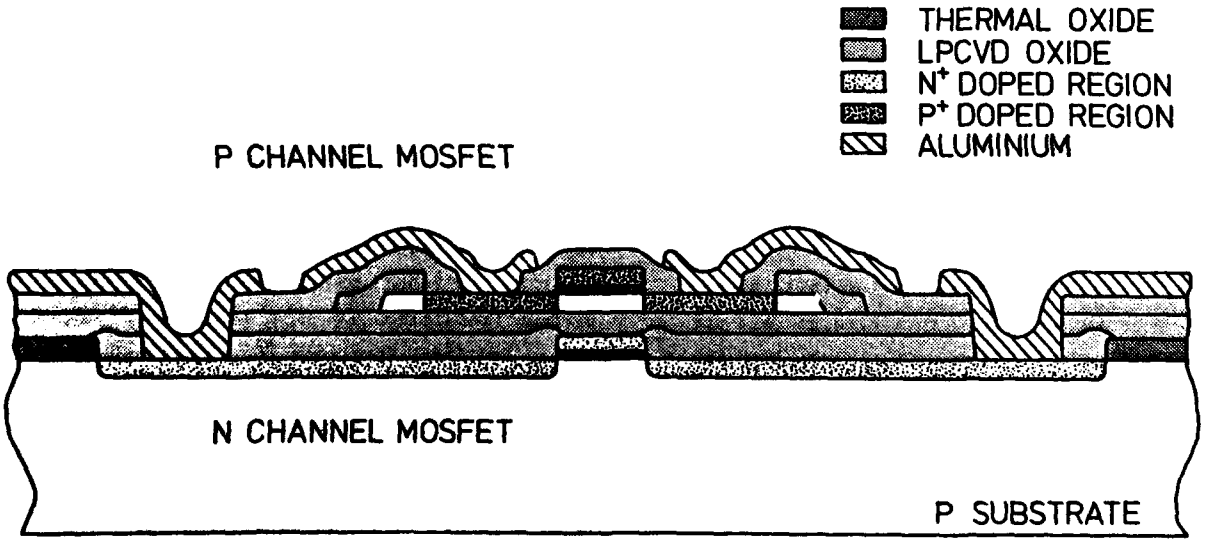


Fig. 2: Schematic cross section of a 3D MOS device.

3. RESULTS

A schematic cross-section of a 3D device consisting of two MOS transistors is shown in Figure 2. A SEM photograph of the corresponding fabricated structure is shown in Figure 3.



Fig. 3: SEM photo of a 3D device.

As stated earlier the main goal of this investigation was a determination of the effect of the recrystallization procedure on the underlying devices in the monocrystalline substrate. Characteristic curves typical of the MOS devices in both layers are shown in Figures 4 and 5. Although the threshold voltage of the n-channel devices is slightly negative this may be easily compensated with an appropriate channel implantation. The p-channel transistors exhibit a relatively high threshold voltage of -12 V and a mobility of roughly 140 cm^2/Vs. This may be attributed to high levels of surface roughness on the boundary layer at the silicon/gate-oxide interface as has been seen with SEM investigation. The source of this roughness may lie in the low scan speed used during recrystallization.

The measurements show that the recrystallization process has a limited effect on the parameters of the devices in the substrate layer. Although the process has no significant effect on the mean values of the threshold voltage and mobility, the parameter scattering increases by roughly 50%. After completion of the second active layer the mean parameter values have shifted as well. It is possible that this shift is caused by the resultant mechanical stress between the

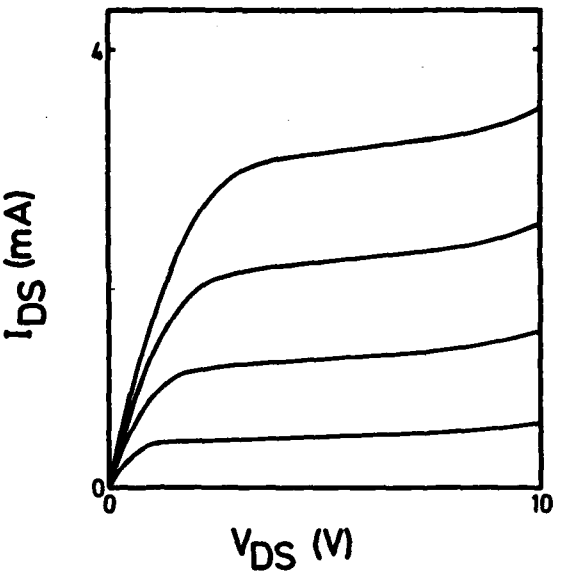


Fig. 4: Typical characteristic curve of a n-channel MOS transistor.

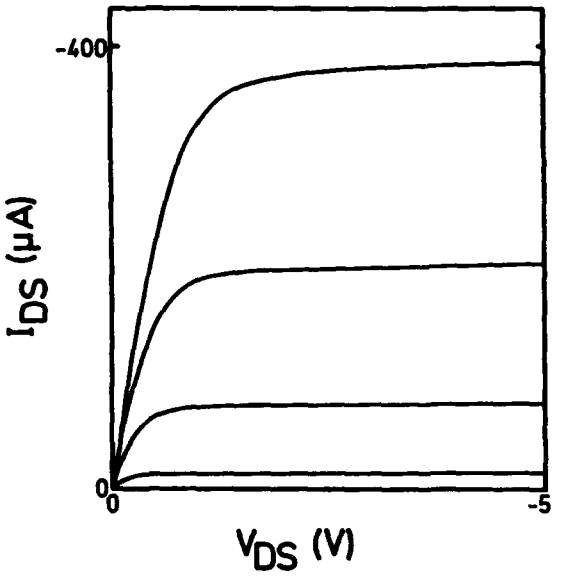
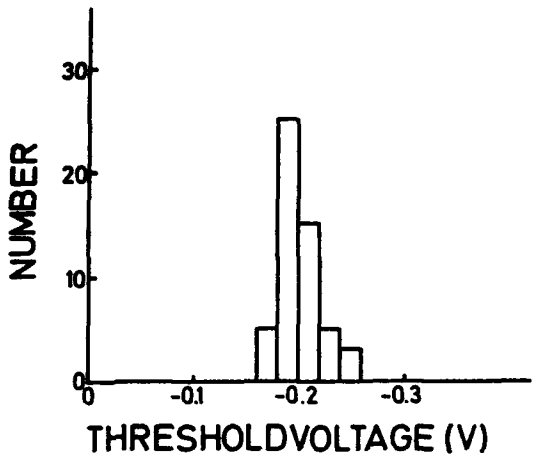
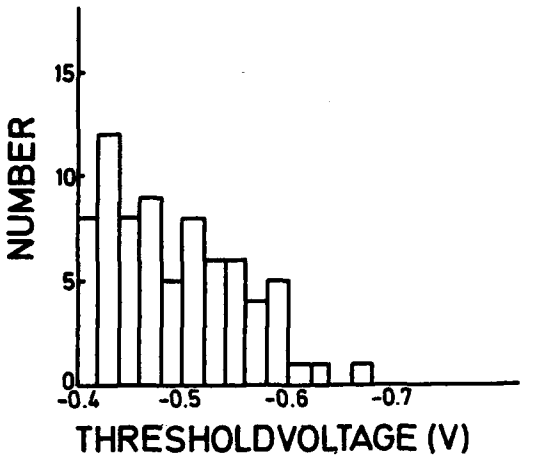
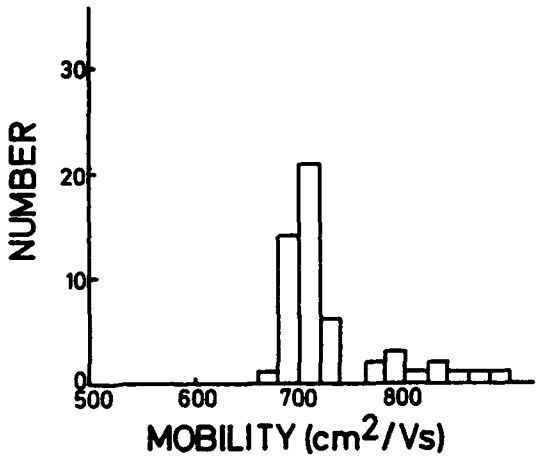


Fig. 5: Typical characteristic curve of a p-channel MOS transistor.







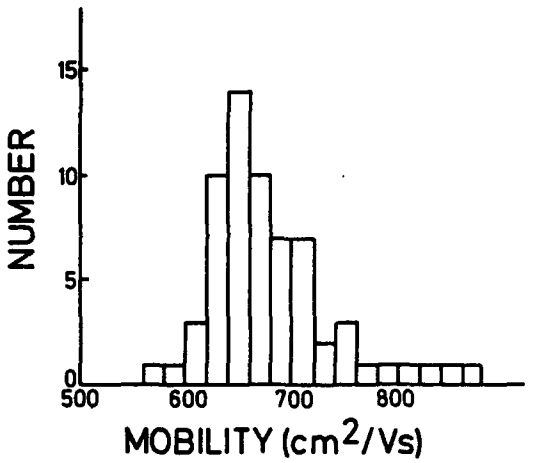


Fig. 6: Measurement distribution of the n-channel transistors after completion of the first layer.

Fig. 7: Measurement distribution of the n-channel transistors after completion of the second layer.

individual layers. The histograms showing these effects are displayed in Figs. 6 and 7.

4. SUMMARY

This investigation has shown that the fabrication of a second active device layer has only a minor effect on the characteristics of the devices in the underlying layer. The process presented here has served as the basis for the development of a 2 um design rule 3D-CMOS process.

ACKNOWLEDGEMENTS

The authors are grateful for the technical support of M. Forster, A. Heidenreich and B. Brandt. This work was supported by the West German Ministry of Research and Development.

REFERENCES

[1] Miyao, M., Ohkara, M., Takemoto, I., Tamara, M., and Tokuyama, T., Appl. Phys. Lett., 41(1), July 1982

VOLUME INVERSION IN SOI MOSFETs WITH DOUBLE GATE CONTROL: A NEW TRANSISTOR OPERATION WITH GREATLY ENHANCED PERFORMANCE

F. BALESTRA, S. CRISTOLOVEANU, M. BENACHIR, J. BRINI and T. ELEWA

Laboratoire de Physique des Composants à Semiconducteurs (UA-CNRS),
ENSERG/INPG, 23 Av. des Martyrs, 38031 Grenoble, FRANCE

Silicon-On-Insulator transistors are used with a double gate control. By this way, a fully inverted silicon film (interface and film volume) is obtained. This method allows us to greatly enhance the device performance, in particular the subthreshold swing, transconductance and drain current. Simulated and experimental characteristics on SIMOX structures are analysed to study the new device.

1. INTRODUCTION

Silicon-On-Insulator materials present many advantages compared with the bulk silicon VLSI technology: lateral isolation, lower parasitic capacitance and power, higher speed, reduced short channel effects, radiation tolerance, ...

In this communication, the special multi-interface configuration of SOI structures is used to obtain a new device based on volume inversion. The theoretical analysis is achieved with a "home-made" computer program (ISIS) which gives the solution of the Poisson equation in multilayer structures [1], and the experiment is carried out on SIMOX devices.

2. SIMULATION

The physical principle of the device is shown in Fig.1. Surface inversion channels can be activated either at the top interface or at the back interface using the normal gate V_{G_1} or the secondary gate V_{G_2} (bulk Si substrate) respectively. We choose to simultaneously bias both gates ($V_{G_2} = K\,V_{G_1}$), where the coefficient K accounts for the differences in thickness and threshold voltage between gate oxide and buried oxide ($K \approx 10$).

If the film is thick or highly doped, there is no overlap of the two depletion regions and the inversion channels grow almost independently. For example in Fig. 1a, the film is slightly depleted for $V_{G_1} = 2V$ and only a low coupling appears between the two conducting channels.

A different behaviour (Fig.1b), caused by the coupling of the two interfaces, occurs in films with normal thickness (< 0.2 μm) and low doping (a few 10^{15} cm^{-3}). For $V_{G_1} = -1$ V, the whole silicon film is in accumulation. For higher V_{G_1}, the potential increases at the interfaces and in the film volume, from depletion to weak and strong inversion. For $V_{G_1} \geq 0.6$ V, the potential shift exceeds $2\Phi_F$ in every regions and all the film is in strong inversion. We propose to call this new device the "volume-*inversion* MOSFET" (VI-MOSFET).

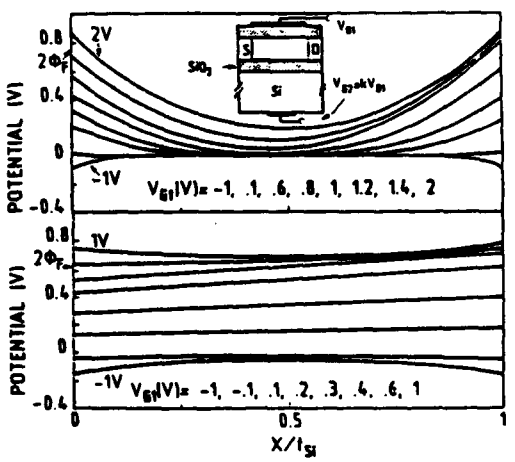


Figure 1

Potential profiles inside the silicon film for uncoupled (a : doping $N_a = 4.10^{16}$ cm^{-3}, film thickness $t_{Si} = 300$ nm) and coupled (b : $N_a = 3.10^{15}$ cm^{-3}, $t_{Si} = 100$ nm) interfaces. (27 nm and 380 nm thick oxides).

The behaviour of the VI-MOSFET is governed by minority carriers, which now are no longer confined at an interface. There are significant advantages: greatly increased number of minority carriers, reduced influence of surface scattering and interface defects, use of the volume which is much thicker than a surface inversion layer and has higher carrier mobility. These special features lead to a great improvement in current value, subthreshold slope, transconductance and speed.

In Fig.2 are compared the current-voltage characteristics of an N-channel VI-MOSFET (K = 10) with those of a normally operated MOSFET (K = 0, i.e. inversion layer at the top surface only). The current and transconductance variation versus V_G are calculated using a mobility profile suggested by transport measurements [2] : 1200 cm^2/Vs in the center of the film, 500 cm^2/Vs at the front interface and 400 cm^2/Vs at the back interface.

The subthreshold swing (Fig.2a) of the VI-MOSFET (29 mV/decade, curve 1) is excellent, compared with 66 mV/decade for the normally operated MOSFET (curve 2), and clearly goes far below the theoretical limit of the normal MOSFET (≈60 mV/decade).

In strong inversion (Fig.2a) the current of the VI-MOSFET (curve 1) exceeds by a factor 3 at $V_{G1} \approx 1.2$ V that of the normal MOSFET (curve 2). This is due to (i) the increase in the total number of carriers, (ii) the improvement of the subthreshold swing and (iii) the transconductance overshoot of the VI-MOSFET (Fig.2b, curve 1). Indeed, the transconductance is clearly enhanced for K = 10 (curve 1) in comparison with K = 0 (curve 2). The maximum field effect mobility (1050 cm^2/Vs), corresponding to the transconductance maximum, is close to the carrier mobility in the film volume.

The comparison with a linear variation of carrier mobility (without peak mobility in the film center) between the two interfaces (curve 3 : K = 10 and curve 4 : K = 0), emphasizes the importance of the volume mobility. It is interesting to note that even for K = 0 (curves 2 and 4) the mobility profile is important, since a volume inversion still exists in a narrow region close to the front surface.

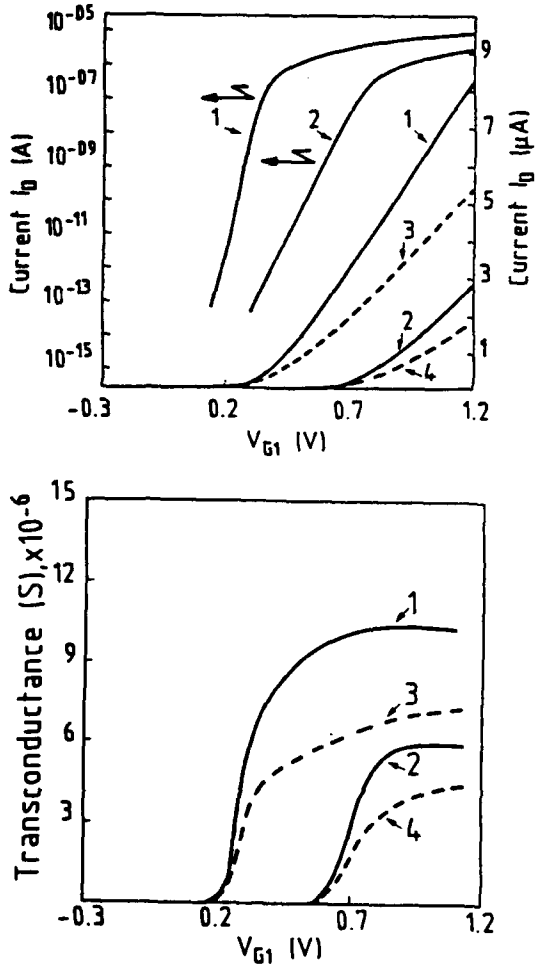


Figure 2

Simulated I-V characteristics (a) and transconductance (b) for K = 10 (curve 1) and K = 0 (curve 2), for the device of Fig.1b and a bell shaped mobility profile. The curves 3 (K = 10) and 4 (K = 0) are obtained with a linear mobility profile. For the sake of simplicity, the mobilities are supposed to not depend on V_G.

Fig. 3 shows the current-voltage characteristics of N-type depletion mode transistor. The drain current and the transconductance are compared for the normally operated MOSFET (K = 0) and for the MOSFET with volume-accumulation (VA-MOSFET) with K = 10. Volume accumulation is more easy to obtain than volume inversion, because it is a "natural" behaviour of the silicon film for a normal transistor in flat band situation. Therefore, the gains are slightly lower for the VA-MOSFET than for the VI-MOSFET. Nevertheless, the use

of a double gate control greatly improves the important device parameters. Indeed, we can observe in Fig.3 a decrease of the subthreshold swing (67 mV/dec for K=0 (curve 2) and 34 mV/dec for K=10 (curve 1)), an increase of the transconductance (90 %) and current value (75 %) for the VA-MOSFET.

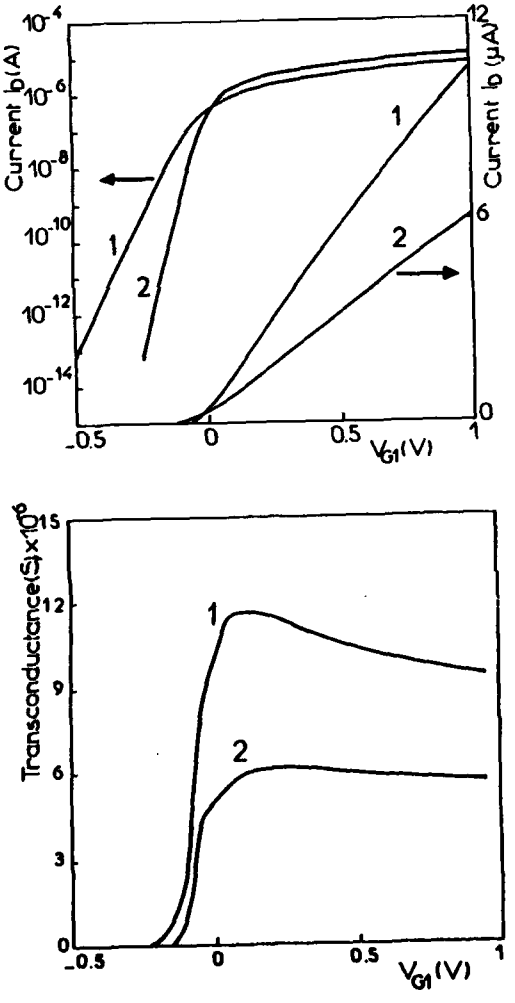


Figure 3

Simulated I-V characteristics (a) and transconductance (b) for K = 10 (curve 1) and K = 0 (curve 2) for a depletion mode transistor with the same technological parameters than in Fig.1b (doping : $N_d = 3.10^{15}$ cm^{-3}).

3. EXPERIMENTS

All these optimistic expectations are indeed verified. Experimental evidence has been obtained with P-channel MOSFETs fabricated with standard technology on a low-doped SIMOX substrate (dose $1.8x10^{18}$ O^+/cm^2, energy 200 keV, annealing above 1300°C). The thicknesses are about 200nm for the film, 27 nm for the gate oxide and 380 nm for the burried oxide. The characteristics of 0.8 μm long transistors are given. Fig.4 a and b clearly show the great gains in subthreshold swing (70 mV/decade for K=0 (curve 1) and 29.5 mV/decade for K=10 (curve 2)), transconductance (80%) and current value which increases by a factor 2.8 at V_{G_2} = -2 V, by using a double gate control. The leakage currents of the VI-MOSFET is very low because carrier accumulation occurs simultaneously at both interfaces. Similar improvements have been obtained on N-channel devices.

The enhancements for the case of depletion mode transistors with volume accumulation are the same than those of the simulated characteristics.

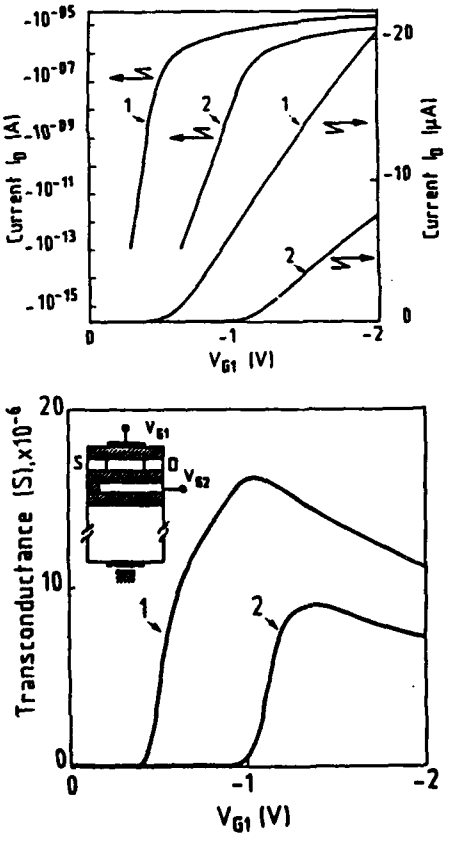


Figure 4

Experimental I-V characteristics (a) and transconductance (b) for a P-channel transistor made on SIMOX.

For practical applications of VI-MOSFETs (or VA-MOSFETs), to overcome the limitation due to the use of the substrate as a unique secondary gate, the double SIMOX structure obtained by two oxygen implants [3] can be proposed (see insert of Fig.4b). By this way, a silicon layer, lying between the two burried oxides, is formed and acts indeed as a gate. Segregation of the secondary gates of various VI-MOSFETs is made by oxidizing the unuseful portion of this film (by simple adjustement of the oxygen implantation) or by etching.

On the other hand, to decrease the voltages applied on the back gate, the difference in thickness between the gate and burried oxide can be reduced. An other solution is to scale down the biases of VLSI circuits.

4. CONCLUSION

In conclusion, the transistor performances have been improved using the new principle of double gate control of VI-MOSFETs or VA-MOSFETs. The experimental enhancements (current, transconductance and subthreshold swing gains) are in agreement with our theoretical calculations.

ACKNOWLEDGMENTS

The authors are indebted to Dr. A.J. Auberton-hervé (LETI, Grenoble) and to Dr. J. Davis (British Telecom, Ipswich) for providing SIMOX devices.

REFERENCES

[1] F. Balestra, J. Brini and P. Gentil, Solid State-Electron. 28, 1031 (1985).

[2] S. Cristoloveanu, S. Gardner, C. Jaussaud, J. Margail, A.J. Auberton-Hervé and M. Bruel, submitted to Journ. Appl. Phys. (1987).

[3] G.K. Celler, J.L. Batstone, K.W. West, P.L.F. Hemment and K.J. Reeson, *IEEE SOS/SOI* Workshop (Captiva, USA), Oct. 1986.

A NEW TYPE OF HIGH PERFORMANCE DEVICE FOR VLSI DIGITAL SYSTEM

XU XIAO-LI, TONG QIN-YI, XONG HE-MING

Microelectronics Center, Nanjing Institute of Technology
Nanjing 210018, China

This paper presents a high performance Complementary Buried Channel FET device isolated by high quality silicon dioxide layer using Silicon wafer Direct Bonding technology (SDB/CBCFET). The structure and operational principle of this device is discussed. The properties of an improved SDB process is investigated. By means of 2D numerical simulation, effects of interface charge density of bonding interface and SOI layer-SiO2 interface on threshold voltage and the threshold voltage shift in submicron geometry are analysed. The performance of submicron SDB/CBCFET device and circuits is evaluated. The results indicate that SDB/CBCFET device is superior to bulk CMOS and SOI/CMOS in speed, switching energy, complexity, reliability and small size effects as device size decreases into submicron dimension.

1. INTRODUCTION

The dominant technology used in modern VLSI digital system is bulk CMOS. The development of bulk CMOS has been supported primarily by rapidly decreasing feature size in the circuits. However, as bulk CMOS technology advances into submicron dimension, several limitations have become apparent: complex process sequences, significant short channel effects and serious interaction between neiboring devices. These negative effects have limited CMOS to improve its performance by further scalling. Silicon-On-Insulator CMOS technique offers an attractive alternative by providing simple device fabrication sequences, improved short channel effects and no latch up problems. However, the poor quality of SOI substrate by conventional techniques such as SIMOX, laser anealling etc. and SOI versions of corresponding bulk CMOS device have limited better use of SOI potential advantages. Moreover, the negative effects of CMOS device can not be effectively suppressed. In recent report (1), we investigated a high quality SOI substrate technology (SDB) which shows that the quality of SOI layer and the underlying SiO2 layer is not degraded from its original "bulk" quality. Based on the characteristics of this technology, we propose a new complementary buried channel device which has higher bulk mobility, much smaller short channel effects and higher performance than bulk CMOS and SOI/CMOS.

In this paper, the structure and operational principle of SDB/CBCFET is described, its small size effects and its device/circuits performance are discussed by means of 2D numerical simulation and analysis models.

2. DEVICE STRUCTURE

SDB/CBCFET is characterized as a compound structure of MESFET and MOSFET (see Fig.1). In its lateral direction, SDB/CBCFET has the same doping type for source, gate and drain which is similar to MESFET, while in its vertical direction, SDB/CBCFET employes poly-Si and SiO2 layer as its gate which is identical to that of MOSFET. By adjusting doping concentration of both poly-Si

gate and channel region, as well as selecting proper SiO2 thickness, the p channel and the n channel device can both be held in deep depletion at zero gate bias and normally off characteristics both for n channel and p channel device can be obtained. Therefore, an n channel device and a p channel device of this type structure can form a basic complementary inverter.(see Fig.

2.1. Properties of SDB Substrate

An SDB technology has been developed. The bonding process adopted in our work involves the following main process sequences: the two mirror polished wafers oxided in wet oxyen, the wafers treated in acid solution, the wafers contacted face to face put into a N2 ambient at 1050C° for about 1 hour forming the SOI substrate. The microstructure and electrical properties of the SOI substrate using this SDB technology have been extensively studied by TEM and DLTS analysis. The experiment results show that dislocations are concentrated on the back side of the substrate and no additional defects have been developed within 80um from SiO2-SiO2 bonding area. Therefore the interface charge density between SOI layer and underlying SiO2 is no more than the interface charge density between conventional bulk silicon and thermal oxidation layer. Both hole and electron mobilities are measured using Van Der Paw method. The results show that hole and electron mobility of the SOI substrate are 369cm /vs and 1079cm /vs respectively which is almost identical to that in original bulk silicon. Since the quality of SOI substrate by SDB technology is not degraded from its original "bulk" silicon quality, this SOI substrate is very suitable for CBCFET device structure.

2.2. Features of CBCFET

In contrast to conventional CMOS, SDB/CBCFET has no p-n junctions at source and drain region. Thus, a significant reduction of small size effects is achieved resulting in much less threshold voltage shift, improved punchthrough resistance, and higher performance.

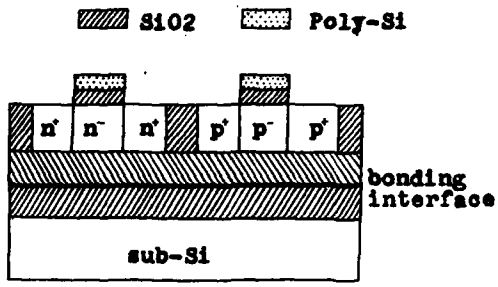


Fig.1 : Cross-Sectional view of SDB/CBC FET configuration

Channel carrier mobility in SDB/CBCFET is high due to its buried channel nature. Furthermore, the low channel doping concentration and the low threshold voltage is responsible for high speed and low power consumption of SDB/CBCFET due to its low logical voltage swing, low supply voltage and high bulk mobility.

3. DEVICE CHARACTERISTICS

The 2D numerical simulation results show that the interface characteristics of SOI layer-underlying SiO2 interface deeply affect device properties. If the interface charge density is large and the thickness of SOI layer is thin, for the n channel device, normally-off operation changes into normally-on operation and for the p channel device, it is in deep depletion status which increases its threshold voltage greatly. Fig.2A shows threshold voltage vs. SOI layer-underlying SiO2 interface charge density for n channel device with various SOI layer thickness. Fig.2B shows the relation between n channel threshold voltage

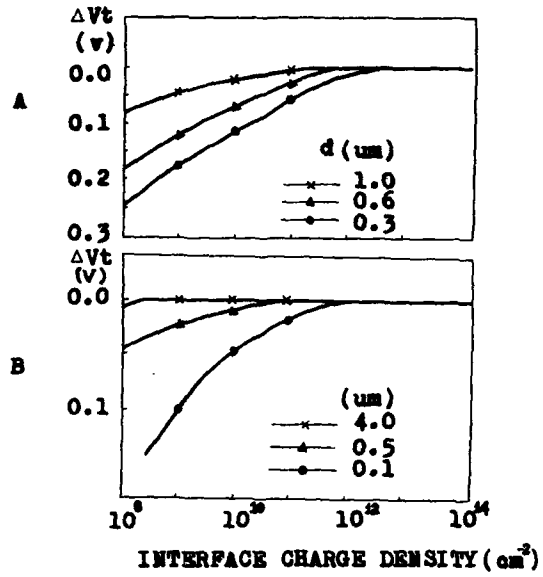


Fig.2
N channel threshold voltage shift vs. interface charge density with 0.5um channel length, 0.78V flat band voltage and 1V supply voltage for A: SOI layer-underlying SiO2 interface with various SOI thickness, B: SiO2-SiO2 bonding interface with various SiO2 thickness.

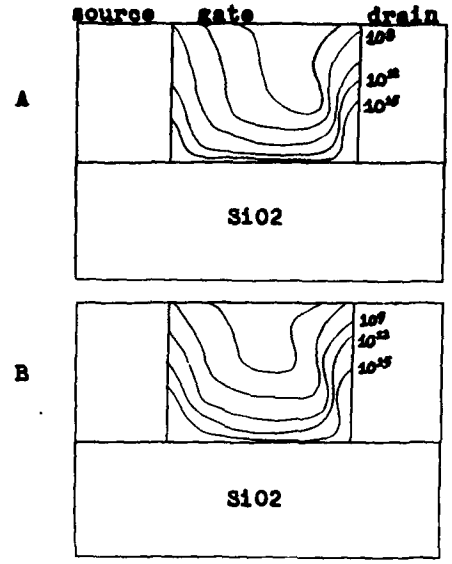


Fig.3 : 2D doping concentration distrabution of n channel device (at o.3V threshold voltage) (A) and p channel device (at -0.3V threshold voltage) (B) with 0.5 um channel length, 0.78V flat band voltage, 5x10[illegible] interface charge density and 1.0V supply voltage.

and bonding interface charge density for various underlying SiO2 layer thickness. From Fig.2B we can conclude that as long as the thickness of underlying SiO2 is larger than 0.5um, the bonding interface charge density will almost not affect the threshold voltage of SDB/CBCFET.

Fig.3A&B shows 2D doping concentration distrabution of n channel device (at 0.3 V threshold voltage) and p channel device (at -0.3V threshold voltage). It is found that both devices show normally-off operation and small punchthrough effects.

Threshold voltage shift as a function of effective channel length for n channel device is shown in Fig.4. The amount of threshold voltage shift of SDB/CBCFET determined by 2D numerical simulation is compared with that of bulk CMOS and that of SOI/CMOS (2). It is shown that for the submicron device, the threshold voltage

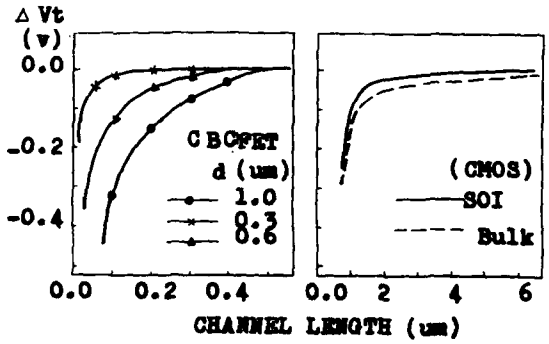


Fig.4 : Threshold voltage shift of n channel device vs. channel length, with 0.78V flat band voltage, 5x10[illegible] interface charge density and 1.0V supply voltage.

shift of n channel device of SDB/CBCFET is much smaller than that of CMOS device structure. The 2D numerical analysis explains the physical origin of the difference between bulk CMOS, SOI/CMOS and SDB/CBCFET. For SDB/CBCFET, absence of

TABLE 1

	L (um)	d (um)	Nd (cm^{-3})	un/up (cm^2/vs)	Vtn/Vtp (v)	Vd (v)	T (ps)	E (fJ)
Bulk CMOS	0.65	5	$1x10^{16}$	453/212	0.7/-0.5	5	460	> 100
SOI/CMOS	0.8	0.5	$1x10^{16}$	680/280	0.4/-0.4	5	93	> 50
SDB/CBCFET	0.5	0.3	$1x10^{16}$	1079/369	0.3/-0.3	1	5.5	3.96

Where L: channel length; d: thickness of Si layer; Nd: doping concentration of active region; un&up: mobility of n channel & p channel device; Vtn&Vtp: threshold voltage of n channel & p channel device; Vd: supply voltage; T: delay time per gate of ring osillator; E: switching energy.

p-n junction at source and drain, small channel depletion charge due to low channel doping and thin SOI layer contribute to much less threshold voltage shift.

4. CIRCUIT PERFORMANCE

In order to show the important characteristics of SDB/CBCFET device and circuits, we have designed a ring osillator composed of SDB/CBCFET device, optimised design parameters and evaluated ring osillator performance (see Table 1). The results indicate that the SDB/CBCFET is superior to bulk CMOS (3) and SOI/CMOS (4) in speed, switching energy and power dissipation as device size advances into submicron dimension.

5. CONCLUSION

A new high performance complementary buried channel device isolated by high quality SiO2 layer using SDB technology has been presented. The characteristics of the device and performance of the circuits have been evaluated. Following results are obtained: 1) simple process sequences, high carrier mobility and small interface charge density of SOI layer by SDB technology, 2) small threshold voltage shift for both n channel and p channel device, 3) high speed, low power, high reliability and complexity of this complementary device. The results indicate that SDB/CBCFET is a potential candidate for very high performance VLSI digital system.

REFERENCES

(1) Li Hui, Tong Qin-Yi, etal., INFOS 87, C.4.4.

(2) Michael P.Brassington, et al., IEEE Trans. Electron Device, vol. ED-32, p1858,1979.

(3) Theodore I. Kamins et al., IEEE Trans. Electron Device Letters, vol. EDL-6, p617, 1985.

(4) A.J.Auberton, et al., IEDM 84, 34.5.

Session C2.2

MOS Modelling
I

Chairman: G. Baccarani

Tuesday, September 15, 1987

COMPARISON OF LONG– AND SHORT–CHANNEL MOSFET'S CARRIED OUT BY 3D–MINIMOS

M. Thurner, S. Selberherr

Institut für Allgemeine Elektrotechnik und Elektronik
Technische Universität Wien
Gusshausstraße 27–29, A–1040 Wien, AUSTRIA

An accurate three–dimensional simulation program for MOSFET devices has been developed by extending MINIMOS (vers. 4) in 3D. The physical model is based on the 'hot-electron-transport model', which includes the Poisson equation, the continuity equations and a selfconsistent set of equations for the currents, mobilities and carrier temperatures. The standard finite difference discretization and the SOR (successive over relaxation) method are utilized to reduce computational time and memory requirements. Adaptive grid refinement is used to equidistribute the discretization errors. Three–dimensional effects like threshold shift for small channel devices, channel narrowing and the accumulation of carriers at the channel edge have been successfully modeled. Our comparison of several MOSFET's make clear that three–dimensional calculations are most important for accurate device modeling.

1 Introduction

The shrinking dimensions of the elements of IC's require for accurate simulation suitable device models in physics and mathematics. The two–dimensional device simulations performed in earlier times described the electrical characteristics for large transistors well but the advanced VLSI technology led to serious problems in modeling such devices and therefore a great demand appeared for 3D simulations.

The three–dimensional effects in MOSFETs like the increasing threshold voltage and the shift of the breakdown voltage caused by the finite channel width are not taken into account by the two–dimensional simulations [1]; the 2D programs are meanwhile state of the art. Accurate investigations of the previously stated effects and the knowledge of increased current densities under certain bias conditions at the channel edge are important not only for studying the electrical device characteristics but also for aging effects [2]–[3]. Therefore we have extended the two–dimensional MINIMOS to a three–dimensional simulation program. A realistic physical model and suitable mathematical algorithms have been developed to simulate the previously stated three–dimensional effects.

We shall report in Chapter 2 about the physics and the mathematics on which the simulations are based.

The results of our simulations carried out by 3D MINIMOS are reported in Chapter 3 and will be discussed there, too. We shall show that the three–dimensional simulations are indispensable for the advance from VLSI– to ULSI technology.

2 The Physical Model and the Mathematical Algorithms for the Three–Dimensional Simulation

The physical model for the simulation program is given by the Poisson and the continuity equations and the drift–diffusions model for the carrier current densities.

$$div\ grad\ \psi = \frac{q}{\varepsilon}(n - p - C) \tag{1}$$

$$div\ J_n = qR \tag{2}$$

$$div\ J_p = -qR \tag{3}$$

$$J_n = -q\mu_n(n\ grad\ \psi - grad\ (U_{t_n} n)) \tag{4}$$

$$J_p = -q\mu_p(p\ grad\ \psi + grad\ (U_{t_p} p)) \tag{5}$$

The Poisson equation (1) will always be solved fully

three–dimensionally; the continuity equations (2) and (3) at the first level of sophistication are solved two–dimensionally in the middle of the channel width. The carrier distribution in the whole volume will be calculated by the assumption of negligible current flow in the third direction $J_{n_z} = J_{p_z} = 0$. Assuming the validity of Boltzmann statistics the previous statement is equal to constant quasi Fermi levels in the direction of the channel width

$$\frac{\partial \varphi_n}{\partial z} = \frac{\partial \varphi_p}{\partial z} = 0$$

So we can write

$$n_{x,y,z} = n_{x,y,\frac{w}{2}} \cdot exp(-\frac{1}{U_t} \cdot (\psi_{x,y,\frac{w}{2}} - \psi_{x,y,z})) \quad (6)$$

$$p_{x,y,z} = p_{x,y,\frac{w}{2}} \cdot exp(+\frac{1}{U_t} \cdot (\psi_{x,y,\frac{w}{2}} - \psi_{x,y,z})) \quad (7)$$

The index $\frac{w}{2}$ denotes the middle of the channel width.

The second level of sophistication is obtained by assuming negligible current flow in the third dimension for the majorities and solving the continuity equation for the minorities fully three–dimensionally.

The third level is the fully three–dimensional solution of the continuity equations for both the minorities and the majorities.

For solving the previously specified set of equations we apply for discretization the standard finite difference method. The grid generation will be performed by an automatic mesh refinement algorithm which equidistributes the discretization error.

The linearized equations are essentially solved with an iterative algorithm. In our case we apply the SOR (Successive Over Relaxation) method. The general iterative algorithm:

$$B \cdot x^{(n+1)} = (B - A) \cdot x^{(n)} + b$$

is solved with the matrix $B = (\frac{1}{\omega} \cdot D - L)$. D is the diagonal part of A which is transformed to the unity matrix while L is the lower triangular part of A. With respect to the special linearization method one unknown reduces to:

$$\begin{aligned} x_i^{(n+1)} = (1-\omega) \cdot x_i^{(n)} + \omega(b_i - \\ - x_{i-1}^{(n+1)} \cdot a_{i-1} - x_{i-NX}^{(n+1)} \cdot a_{i-NX} - \\ - x_{i-NXY}^{(n+1)} \cdot a_{i-NXY} - \\ - x_{i+1}^{(n)} \cdot a_{i+1} - x_{i+NX}^{(n)} \cdot a_{i+NX} - \\ - x_{i+NXY}^{(n)} \cdot a_{i+NXY}) \end{aligned} \quad (8)$$

in which $i = 1...NX \cdot NY \cdot NZ$ (NX points in $x-$direction, NY points in $y-$direction and NZ points in $z-$direction).

The advantage of this method is given by the small amount of memory requirement, precondition work and relatively fast convergency,as well. Through an adaptive determination algorithm for the optimum relaxation factor ω we use only a moderate amount of CPU time [5]. The system of the coupled nonlinear difference equations are solved with Gummel's iterative method.

3 The Numerical Results and Discussion

With the previously given physical model a three–dimensional MOSFET simulation program has been developed. We have investigated several MOSFETs with this program, two of the investigated devices are presented in comparison and the results discussed.

Both investigated devices are of the same geometrical shape and dimensions (Fig. 1 and Fig. 2) except the channel lengths which are $5\mu m$ and $1\mu m$ for of device 1 and 2,respectively. The channel widths are $1\mu m$,the gate oxide thickness $15nm$, the substrate doping $2 \cdot 10^{16} cm^{-3}$ and the source/drain doping $1.69 \cdot 10^{20} cm^{-3}$. In Fig. 1 and Fig. 2 the field oxide which limites the channel in the third dimension, can be seen at the backside of the MOS model. The contacts of source and drain which are left and right in the figures 1 and 2, extend over the channel width, whereas the gate contact covers the channel and the field oxide. The shape of the field oxide in our case is approximated by a rectangular geometry.

The potential distribution at the bias condition $U_{DS} = 2.0V$, $U_{BS} = 0.0V$ and $U_{GS} = 3.0V$ can be seen in Fig. 3 and Fig. 4. The threshold voltages

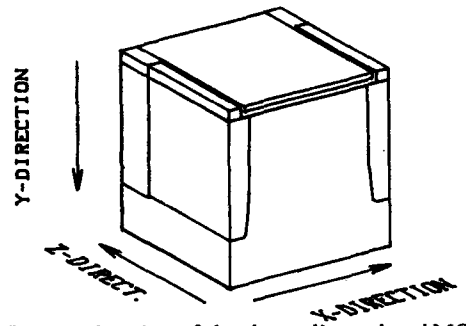


Fig.1: Perspective view of the three–dimensional MOS-FET structure with channel length of $5\mu m$ and channel width of $1\mu m$.

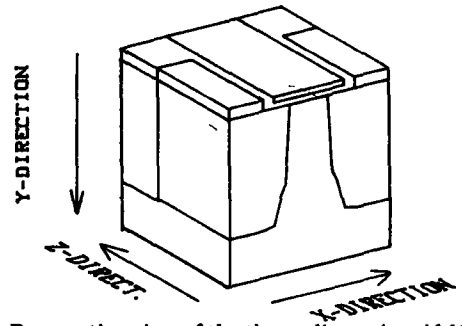


Fig.2: Perspective view of the three–dimensional MOS-FET structure with channel length of $1\mu m$ and channel width of $1\mu m$.

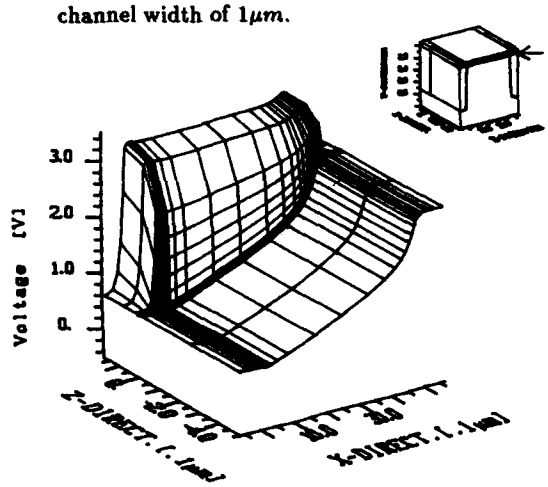


Fig.3: 3D–plot showing a detailed view of the surface potential at the channel edge along the channel length for device 1 at bias $U_{DS} = 2.0V, U_{BS} = 0.0V, U_{GS} = 3.0V$.

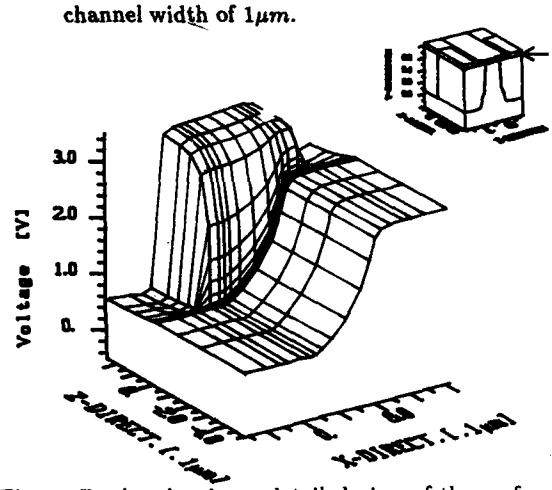


Fig.4: 3D–plot showing a detailed view of the surface potential at the channel edge along the channel length for device 2 at bias $U_{DS} = 2.0V, U_{BS} = 0.0V, U_{GS} = 3.0V$.

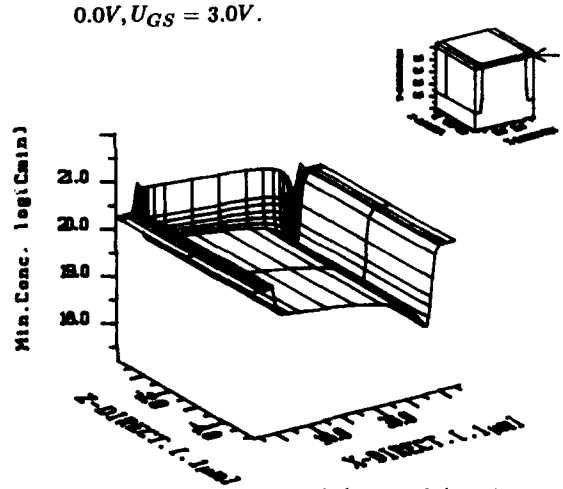


Fig.5: 3D–plot showing a detailed view of the minority density at the channel edge along the channel length for device 1 at bias $U_{DS} = 2.0V, U_{BS} = 0.0V, U_{GS} = 3.0V$.

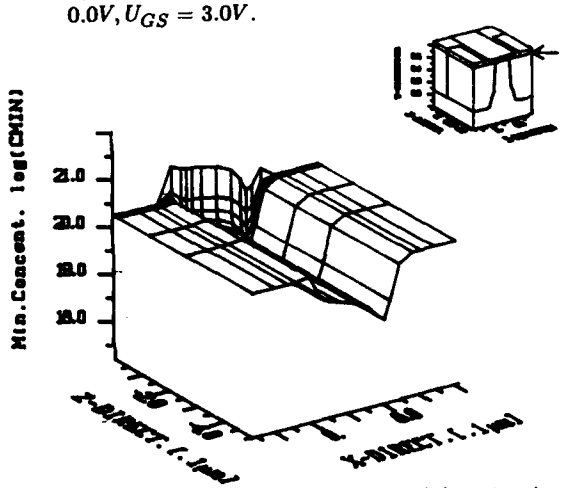


Fig.6: 3D–plot showing a detailed view of the minority density at the channel edge along the channel length for device 2 at bias $U_{DS} = 2.0V, U_{BS} = 0.0V, U_{GS} = 3.0V$.

are $U_{th} = 0.74V$ and $U_{th} = 0.62V$ for device 1 and 2, respectively. That means that we are far above threshold. Note the strong increase of the potential in the field oxide. The minority (electron) distributions (Fig. 5 and Fig. 6), show a very interesting effect. The carrier densities in the short channel MOSFET (device 2) is much higher compared to that of the long channel MOSFET (device 1). The accumulation of the minorities at the channel edge at the given bias condition is based on the limitation of the channel width and the high potential in the oxide region. In the subthreshold region this effect will change into its opposite. The currents calculated by two–dimensional simulations at the previously described bias conditions will be much smaller compared to that of three–dimensional simulations. These effects increase with shrinking device dimensions. Both short and small channel devices will be very sensitive on geometrical changes in respect to their electrical characteristics. Especially we expect that the increase of minorities at the channel edge influences the reliability and safety of devices.

Not only the previously discussed effects but also other three–dimensional effects like the shift of the threshold voltage and the increased breakdown voltage at small channel MOSFETs have been simulated with our program. All these effects are known from theory and from practical measurement but not satisfactorily modeled until now.

Acknowledgement

This work was supported by the research laboratories of SIEMENS AG at Munich, FRG, by DIGITAL EQUIPMENT CORP. at Hudson, USA, and by the "Fonds zur Förderung der wissenschaftlichen Forschung", project S43/10. We are indepted to Prof. H. Pötzl for many helpful discussions.

References:

[1] S.M. Sze, Physics of semiconductor devices, ISBN 0-471-09837-X, John Wiley & sons, 1981

[2] S.M. Sze, VLSI–Technology, ISBN 0-07-062686-3, McGraw-Hill, 1983

[3] T. Iizuka,K.Y. Chiu, and J.L. Moll, Double threshold MOSFETs in bird's–beak free structures, IEEE Int. Electron Device Meet.,Wash.,D.C., 1981, p. 380

[4] L.A. Hageman,Franklin T. Luk,David M. Young, On the equivalence of certain iterative acceleration methods, SIAM J.NUMER.ANAL., pp 852-873, vol. 17 No. 6, Dec 1980

[5] R.G. Grimes,D.R. Kincaid,D.M. Young, ITPACK 2A – A fortran implementation of adaptive accelerated iterative methods for solving large sparse linear systems, Report CNA-164, Center for numerical analysis, University of texas at Austin, 1980

[6] S. Selberherr, The status of MINIMOS, Proc. Simulation of semiconductor devices and processes, pp 2-15, Swansea, 1986

[7] S. Selberherr, Analysis and simulation of semiconductor devices , ISBN 3-211-81800-6, Springer, WIEN NEW-YORK, 1984

[8] O. Axelsson, Solution of linear systems of equations;Iterative methods, Lecture notes in mathematics 574, SMT, 1976

THREE-DIMENSIONAL SIMULATION OF A NARROW-WIDTH MOSFET

P. Ciampolini, A. Gnudi, R. Guerrieri, M. Rudan and G. Baccarani

Dipartimento di Elettronica, Università di Bologna
viale Risorgimento 2, 40136 Bologna, Italy

Abstract

In this paper we illustrate the main features of a general-purpose three-dimensional device-analysis program, HFIELDS-3D, developed at the University of Bologna in the context of an EEC-supported ESPRIT Project. The program employs triangular-based prismatic elements, which provide a reasonable compromise between simplicity and flexibility, but it is not otherwise limited to any specific device structure. In the present implementation, the program handles Poisson and one-carrier continuity equation, which allows for the simulation of unipolar devices. As an example, a typical 3-D problem, the narrow-width effect, is investigated using a realistic device structure fully accounting for the typical bird's beak. It is shown that not only the threshold voltage, but also the gain factor, and therefore the device transconductance, are affected by the narrow-width effect.

1. Introduction

Numerical simulation of semiconductor devices in two dimensions is nowadays a well-established technique for the design of advanced electronic components and processes. As device miniaturization progresses toward submicron feature sizes, however, 3-D effects are getting more and more important even for nominally-standard planar devices, thus making two-dimensional simulation codes inadequate for predicting device performance. In addition, increasingly complex device geometries are being devised, such as the buried-electrode dynamic RAM cell, currently used in high-capacity memory devices, the floating-gate EPROM cell, and the I^2L NOR gate, which are inherently three dimensional. All the above devices can only be simulated by means of 3-D device-analysis programs.

Most of the activity reported so far in this field has been performed in Japan [1-3] and, to a lesser extent, in the United States [4]. In this paper we illustrate the main features of a general-purpose three-dimensional code, HFIELDS-3D, developed at the University of Bologna in the context of an EEC-supported ESPRIT Project, and show how such a code can be profitably used to investigate a classical three-dimensional problem, i.e. the narrow-width effect in MOSFET's.

The program employs triangular-based prismatic elements, which allow for a reasonable compromise between geometrical flexibility and simplicity of implementation, but it is by no means restricted to any specific device structure. In the present stage of development, only Poisson and one carrier-continuity equation are solved, thus making the program suitable for the simulation of unipolar devices.

The next section discusses the fundamental choices of the present project, and provides some details on the software implementation of the program. Numerical results are illustrated in section 3 and conclusions are drawn in section 4.

2. Features of the program

HFIELDS-3D allows up to ten semiconductor and insulator regions, which can be either simply or multiply connected. An equal number of insulator-semiconductor interfaces, ohmic contacts, gates and floating gates are allowed. Thus, rather complex device structures can be accommodated by the program.

As already anticipated, HFIELDS-3D employs a triangular-based prismatic-element mesh which, in the authors' opinion, is flexible enough from the geometrical standpoint for most practical applications, while still having a number of implementation advantages over tetrahedral meshes:

- The problem of properly defining the control volumes associated with the nodes of general tetrahedra is still largely unsolved or, at least, insufficiently tested. As a result, some ripple could be expected in the resulting solution.
- The problem of generating tetrahedral meshes in three dimensions which take real advantage of the potential flexibility of these elements (i.e. without converting prisms into tetrahedra) is a very hard task. On the other hand, generating a prismatic-element mesh can be easily accomplished by simply replicating a triangular mesh in the third dimension. It should be noticed that the above procedure does not imply a geometrical uniformity of the simulated device in the third dimension, but it requires step-like changes.

In its present implementation the program handles Poisson and one carrier continuity equation (electrons), which are solved using the Gummel [5] successive procedure. The adopted discretization scheme is the well-known "box integration method", whereby each node is allocated a control volume which, in our case, is still a prism. Poisson's equation is discretized assuming a piecewise linear approximation for the electric potential along the mesh lines, and a three-dimensional generalization of the Sharfetter-Gummel scheme [6] is used to discretize the current-continuity equations. The discretization procedure of Poisson's equation leads to a set of N non-linear algebraic equations, while the corrent-continuity equation leads to a linear system (if the dependence of the carrier mobility upon carrier concentration is neglected).

For the solution of the linear system, we employ the ICCG method for Poisson's equation, and a direct solver for the current-continuity equation. Parallel techniques are currently being investigated for a subsequent vector-processor implementation of the code.

3. The narrow-width MOSFET

Several authors [7,8] have published experimental and/or theoretical results on narrow-width MOSFET's. From the above papers, one can infer that the most important narrow-width effect is an increase of the threshold voltage as the channel width becomes narrower.

We have simulated the electrical properties of a narrow-width MOSFET having a fixed, nominal channel length $L = 1.0\,\mu m$ and a nominal channel width W ranging from $0.4\,\mu m$ to $2.8\,\mu m$. The mesh of one of the above devices is shown in fig. 1: due to the simmetry of the structure, only half of the device is actually considered. The "bird's beak" at the transition between the channel and the field regions is carefully described on the front plane in order to take full advantage of the flexibility of the triangular mesh; so doing, current flow occurs mainly in the third direction, i.e. normal to the front plane.

The 2-D mesh is replicated in the third dimension, where 19 planes are accommodated, with slight modifications which account for the structural changes of the MOSFET (transitions between the source-gate and gate-drain regions). The whole mesh comprises 2,565 nodes and 4,158 prisms.

The impurity concentration is input via analytical expressions reasonably accounting for source and drain diffusions, channel implant and channel-stop diffusion. The channel implant was designed to give a threshold voltage $V_T = 0.7\,V$.

All simulations were carried out on a MicroVAX-GPX work station, and the average CPU time required was about 1.5 hours per bias point.

Figure 2 shows a perspective plot of the equilibrium potential in the plane normal to the X-axis, located at the field-oxide silicon interface. The two upper "plateau" represent the source and drain regions, and the channel appears as a saddle between them. Figure 3 shows instead the electric potential in the parallel plane located at the gate-oxide silicon interface. The ridge at the periphery of the channel is due to the penetration of the device cross section into the oxide and not to an actual increase of the electric potential at the Si-SiO_2 interface. Rather, the fringing field arising from the field-oxide penetrates in the channel region, causing a decrease of the electric potential at its edges and thus reducing the effective inversion layer width. Consequently, when applying a drain-source voltage, a corresponding decrease in the corrent flowing along the channel is to be found.

In order to highlight this effect, some comparisons with 2D results have been made: a corresponding planar device has been simulated neglecting the field oxide and the bird's beak, and assuming equal channel widths. Figure 4 shows the simulated turn-on characteristics of the 2-D and 3-D MOSFET's for three different values of the channel width, namely $W = 0.4\,\mu m$, $W = 1.2\,\mu m$ and $W = 2.8\,\mu m$, respectively. The drain-source voltage $V_{DS} = 0.1\,V$ in these simulations. The figure shows that the 2-D MOSFET systematically provides a larger current than the 3-D one, and the difference increases as the gate voltage is increased. Thus, both threshold voltage and gain factor are affected by the narrow-width effect. Figure 5 shows the corresponding turn-on characteristics with $V_{DS} = 3.0\,V$, which confirm the above statement. The subthreshold behaviour of the turn-on characteristics is better illustrated in figure 6. For the 0.4 μm-width device, over one order of magnitude difference between the 2-D and 3-D models is observed.

A comprehensive view of the I_D/W dependence upon gate voltage and channel width is given in figures 7 and 8, where a perspective plot of a 2-D surface in the 3-D space is shown. These plots emphasize the transconductance degradation and the change in threshold voltage which occur for small values of the channel width.

4. Conclusions

In this paper we have illustrated the main features of a general-purpose three-dimensional device-analysis program, HFIELDS-3D, developed at the University of Bologna in the context of an EEC-supported ESPRIT Project. Care has been taken to ensure versatility and geometrical flexibility of the code, which allows for a wide variety of realistic device structures. At the present stage of development, the program solves only Poisson's and one carrier-continuity equation, which makes it suitable for the simulation of unipolar devices, but its extension to both carrier-continuity equa-

tions is forseen in the near future. The program has been shown to perform satisfactorily in the investigation of the narrow-width effect in MOSFET's, but much work is still to be done in order to optimise its numerical efficiency. To this purpose, the authors are currently investigating more efficient algorithms for a vector-processor version of the code. In order to handle the huge number of equations inherent in 3-D problems in a reasonable time, the use of vector processors or parallel architectures is mandatory.

Acknowledgements

This work has been partially supported by EEC under the ESPRIT 962E-17 Project.

References

1. T. Toyabe, H. Masuda, Y. Aoki, H. Shukuri and T. Hagiwara: "Three-dimensional device simulator CADDETH with highly convergent matrix solution algorithms", *IEEE Trans. on Electron Devices*, vol. ED-32, pp. 2038-2043, 1985.
2. K. Yokoyama, M. Tomisawa, A. Yoshii and T. Sudo: "Semiconductor Device Simulation at NTT" *IEEE Trans. on Electron Devices*, vol. ED-32, pp. 2008-2017, 1985.
3. N. Shigyo and R. Dang: "Three-Dimensional Device Simulation Using a Mixed Process/Device Simulator", from *Process and Device Modeling*, Ed.: W. L. Engl, North Holland, 1986.
4. E. M. Buturla, P. E. Cottrell, B. M. Grossman, C. T. McMullen and K. A. Salsburg: "Three-Dimensional Transient Finite-Element Analysis of the Semiconductor Transport Equations", from *Numerical Analysis of Semiconductor Devices, Proc. of the NASECODE II Conference*, pp. 160-165, Boole Press, Dublin, 1981.
5. H. K. Gummel: "A Self Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculation", *IEEE Trans. on Electron Dev.*, vol. ED-11, pp. 455-465, 1964.
6. D. L. Sharfetter and H. K. Gummel: "Large-Signal Analysis of a Silicon Read Diode Oscillator", *IEEE trans. on Electron Devices*, vol. ED-16, pp. 64-77, 1969.
7. L. A. Akers, M. M. E. Beguwala and F. Z. Custode: "A Model for a Narrow Width MOSFET Including Tapered Oxide and Doping Encroachment", *IEEE Trans. on Electron Dev.*, vol. ED-28, pp. 1490-1495, 1981.
8. Y. C. Cheng and P. T. Lai: "An Analytical Model for the Threshold Voltage of a Narrow-Width MOSFET", *IEEE Trans. on Electron Dev.*, vol. ED-31, pp. 1814-1823, 1984.

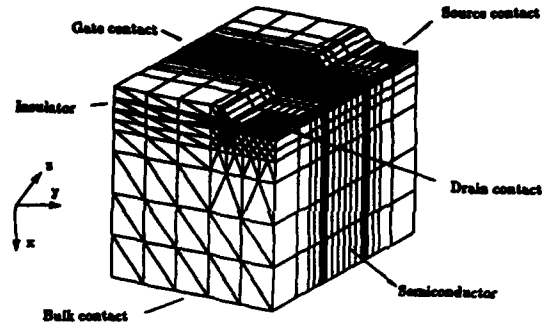


Fig. 1: Mesh of the simulated MOSFET.

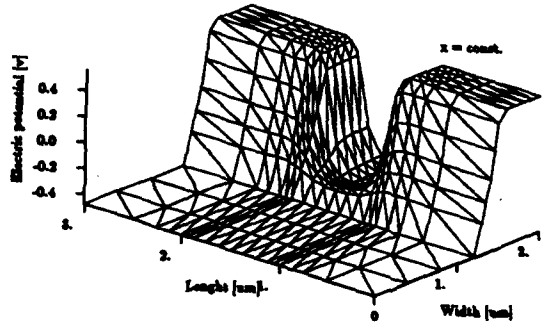


Fig. 2: Perspective plot of the equilibrium potential in the plane normal to the X-axis, located at the field-oxide silicon interface.

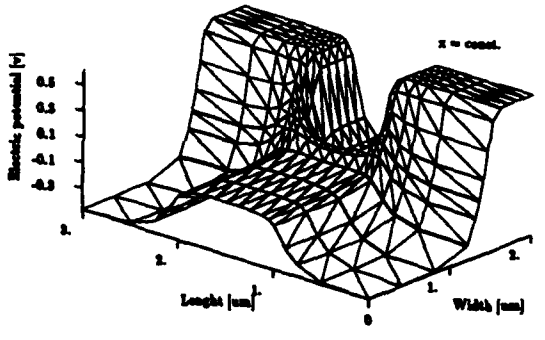


Fig. 3: Perspective plot of the equilibrium potential in the plane normal to the X-axis, located at the gate-oxide silicon interface.

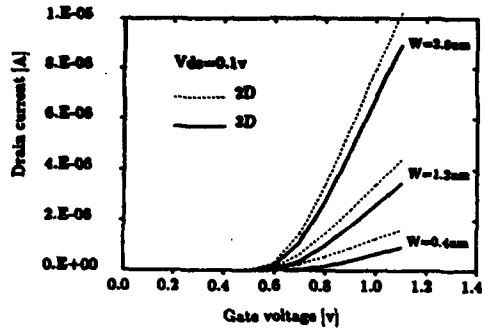


Fig. 4: Simulated turn-on characteristics of the 2-D and 3-D MOSFET's for different values of the channel width; $V_{ds} = 0.1V$.

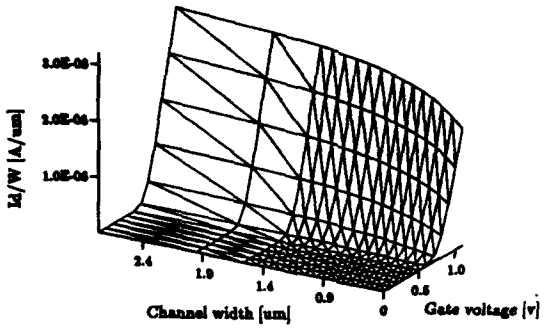


Fig. 7: Normalized drain current as a function of the channel width and the gate voltage.

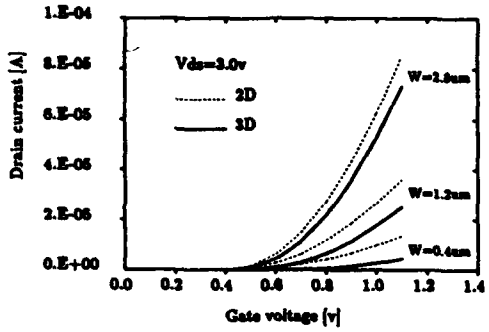


Fig. 5: Simulated turn-on characteristics of the 2-D and 3-D MOSFET's; $V_{ds} = 3.0V$.

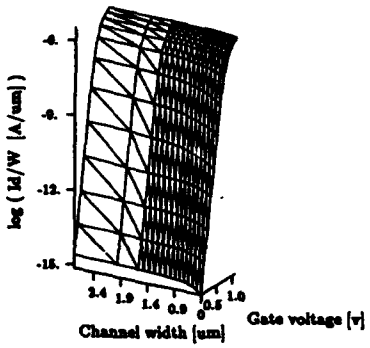


Fig. 8: Normalized drain current as a function of the channel width and the gate voltage.

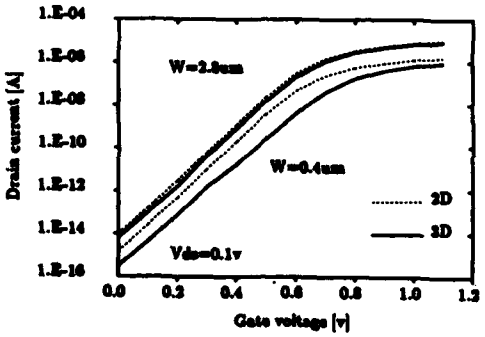


Fig. 6: Simulated turn-on characteristics of the 2-D and 3-D MOSFET's; $V_{ds} = 0.1V$.

2-D AND 3-D CAPACITANCE EFFECTS IN MOS VLSI

J.H.M.M. Quint, F.M. Klaassen, R. Petterson

Philips Research Laboratories
P.O.Box 80000, 5600 JA Eindhoven, The Netherlands

ABSTRACT

Spreading capacitances of several MOS VLSI configurations have been calculated numerically by solving Poisson's equation in 2 or 3 dimensions. Owing to nonuniform charge distributions, contributions from sidewalls and topsurfaces, and shielding effects, considerable deviations from scarce analytic formula have been found. Successively considered are the cases: 3 parallel conductors at equal height from the substrate, 2 parallel conductors at different level from the substrate, gate-drain configuration of different MOSFETs, and two conductors or four conductors crossing above a substrate.

1. Introduction

Since bias voltage constraints and electromigration effects prevent to scale down properly the dimensions of the interconnection system, in submicron VLSI additional capacitance effects become more and more important. Not only do charges at the sidewalls and top surfaces of the conductors lead to a larger capacitance to substrate or to interline capacitance, but nonuniform charge distribution and shielding effects cause the capacitance to deviate considerably from the 1-D value.

Although analytic capacitance expressions based on conformal mapping have been given to correct for several of the above effects [1,2,3], it is questionable whether these results are generally useful. Not only are the results limited owing to several assumptions used, but in practical layouts configurations soon become too complicated to use conformal mapping. Since an empirical investigation is costly and sometimes even impossible, the capacitances from a number of elementary interconnect configurations have been calculated numerically by making use of the device simulator TRIPOS [4], which solves Poisson's equation in two and three dimensions. Where possible, the results have been compared to analytic results. Furthermore the observed trends are discussed from a physical point of view.

2. Capacitances between 3 parallel conductors and the substrate

Owing to 2-D and 3-D charge distribution effects, the capacitance per unit length deviates from the conventional 1-D formula $C = \epsilon_{ox} W/H$, when the distance (H) to substrate and the thickness (T) of conductors are no longer small compared to their width (W) and mutual distance (S). This is shown in fig. 1, where all possible capacitances of 3 conductors parallel to each other and to the substrate are given as a function of the distance S (with $T = W = 1$ μm). Lateral field distributions underlying the above 2-D results are given in fig. 2 (with the location indicated in the inset of fig. 1). While the normal field remains rather uniform under the central conductor owing

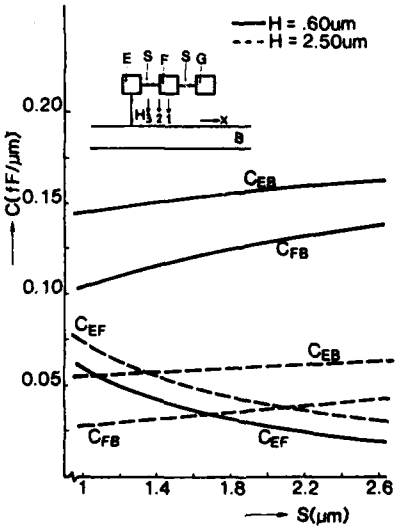


Figure 1. Capacitances vs. mutual distance for conductor configuration given in the inset.

to shielding by the surrounding conductors, the lateral field is very nonuniform. Starting from nearly zero value along the symmetry line (1), the above field peaks to a high value at the edge (2) of the central conductor due to charge crowding at the corners of the conductor (at which upto 30% of the total charge may be accumulated).

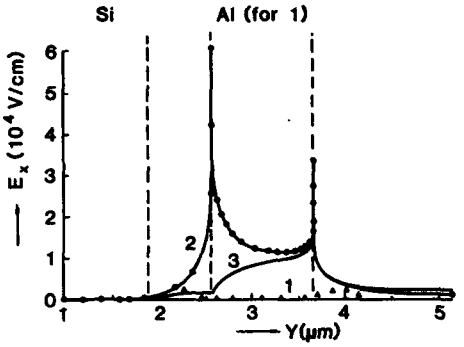


Figure 2. Lateral field along several normal directions for configuration of fig. 1.

Owing to shielding effects of field lines originating from the sidewalls and top surfaces the capacitance to substrate of the central conductor ($C_{FB} = \Delta Q_B/\Delta V_F$) is lower than that of the left conductor and for smaller values of the distance S the difference becomes larger. At the same time the interline capacitance ($C_{FE} = \Delta Q_E/\Delta V_F$) increases with decreasing S and at a higher rate, than the others decrease. As shown in fig. 3 the total capacitance of the middle conductor $C_T = C_{FB} + 2C_{FE}$ increases with decreasing S. In the same figure also a semi-empirical approximation [3] is given. Although the deviation with TRIPOS remains within 10%, the deviation for the various parts of C_T is larger.

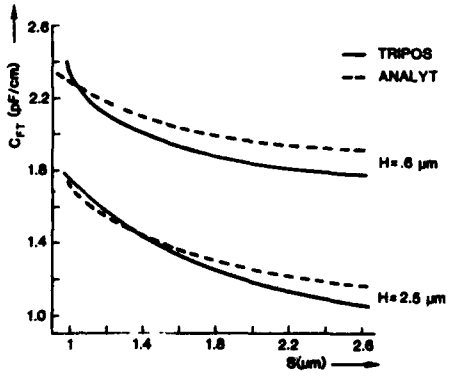


Figure 3. Total capacitance of central conductor of fig. 1.

3. Capacitances between two conductors at different level

When two conductors are parallel, but located at different levels compared to the substrate, their partial overlap has a specific effect on all possible capacitance values. This is shown in fig. 4, which gives the capacitances as a function of the overlap distance S as defined in the inset. Owing to shielding of field lines originating from the top surface of conductor L, the capacitance to substrate C_{LB} at complete overlap ($S = 0$) is lower than the value in the absence of conductor T, and even slightly decreases at small partial overlap ($S < 0.3$ μm). Only when the overlap approaches zero ($S > 1.0$ μm) this capacitance increases with S. Because of differences in the shielding of field lines originating from the bottom and sidewalls of conductor T, for the capacitance C_{TB} the same qualitative behaviour is observed. However, owing to the fact that at $S = 0$ the bottom surface of T has been shielded effectively, the relative effects are smaller. Furthermore, while at complete overlap the interline capacitance C_{TL} varies strongly with the step height ($H_T - H_L$), for larger values of S the above capacitance becomes completely determined by fringing effects and therefore hardly varies.

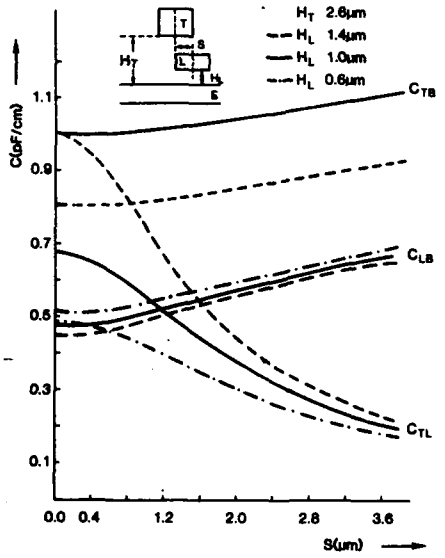


Figure 4. Capacitances vs. mutual distance for configuration given as inset (thickness of $L = .6$ μm).

4. Gate-drain capacitance of a MOSFET

Below threshold voltage the gate-drain capacitance ($C_{GD} = \Delta Q_G/\Delta V_D$) is larger than expected from the physical overlap owing to fringing effects between the poly silicon gate sidewall and the drain topsurface, and

between the drain-channel edge and the gate lower surface (compare the inset of fig. 5). For the conventional MOSFET an analytic expression for the above capacitance

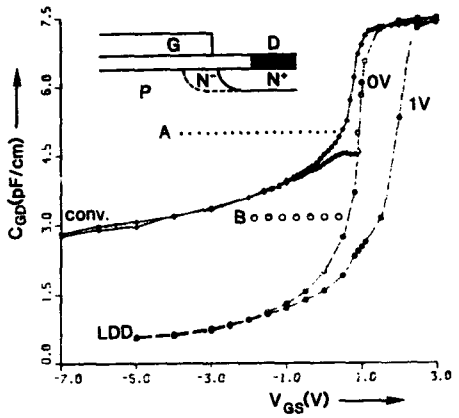


Figure 5. Gate-drain fringing capacitance of conventional and LDD MOSFET (lines (A, B) are analytic results).

has been given [5], but the validity of this approach is questionable for an LDD-MOSFET with its refined gradual drain junction profile. For both devices the numerically calculated value of C_{GD} is given in fig. 5 as a function of the gate voltage at two different values of the drain voltage. The devices considered are n-channel type with a poly silicon gate (width 0.7 μm, thickness 0.3 μm) on a 17 nm thin insulator. For the conventional MOSFET the 0.25 μm drain junction has a physical overlap with the gate of 0.17 μm. Using process modelling data the lightly doped drain of the LDD-type is assumed to overlap the gate to an amount of 0.08 μm.

Decreasing V_G from zero Volt (off-state) causes the building-up of an accumulation charge starting from the centre of the channel region. This growing accumulation charge gradually reduces the inner fringing capacitance between gate and drain edge. In the sub-threshold region C_{GD} increases with V_G owing to the gradual formation of a channel starting from the drain. This generally increases the 1-D overlap. At threshold voltage the capacitance saturates with V_G, since the channel is present everywhere. However, due to channel pinch-off occuring at V_D = 1 Volt C_{GD} remains lower than the value at V_D = 0. Owing to the smaller 1-D overlap and partial depletion of the lightly doped drain region (in particular at V_D = 1 Volt) the capacitance of the LDD MOSFET is considerably smaller in the accumulation and subthreshold region.

Also shown in fig. 5 is an analytic result [5]. Only for the conventional device at V_G = 0 this result fairly agrees with the numerical value. However the reduction of the fringing effect in deep accumulation is not taken into account.

5. Capacitance of two crossing conductors

When a second conductor crosses a first one at different height from the substrate (inset of fig. 6) the field and the resulting charges have to be calculated using a full 3-D solution of the Laplace equation [6]. Fig. 6 gives the three possible capacitance values as a function of distance between both conductors. The upper conductor T and the lower conductor S have a thickness of 1.1 μm and 0.6 μm respectively, a width of 1 μm and a length of 20 μm.

Of course the most pronounced effects are found in the interline capacitance C_{TS}. Not only is the value owing to 3-D charge contributions almost an order higher than the 1-D capacitance value, but the presence of the substrate causes that the above value is still smaller than the pure interline capacitance value C'_{TS} (substrate removed). In addition the contributions from sidewalls and top/bottom areas cause the capacitance C_{TS} to vary sublinearly with H^{-1}. Shielding effects are the main reason for C_{SB} to increase slightly with increasing H.

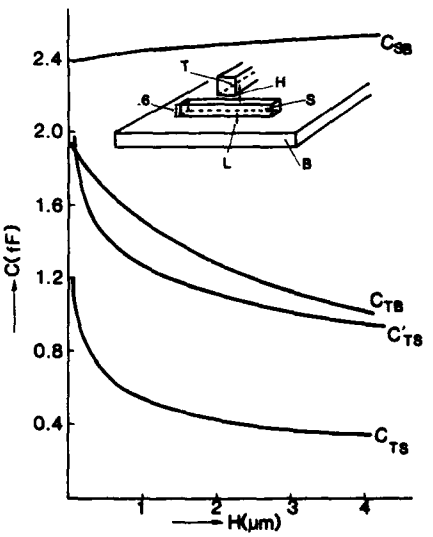


Figure 6. Capacitances vs. distance for two crossing conductors above a substrate (L = 1 μm).

6. Capacitances of four conductors

As an extension to the previous case a configuration of four conductors is considered, which are crossing mutually and are located at two different levels above the

substrate (compare the inset of fig. 7). Naturally the same 3-D effects occur as in the previous case, but owing to the

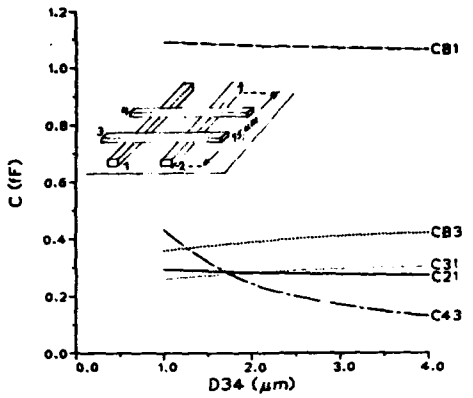


Figure 7. Capacitances between four conductors (see inset) vs. distance of upper conductors ($d_{13} = d_{1B} = 1$ μm, $d_{12} = 2$ μm).

more complicated situation the calculation is more demanding for the simulator. In spite of making use of symmetry rules the number of grid points required amounts 30.000.

In fig. 7 several capacitances to bulk and interline capacitances have been plotted as a function of the distance between the two upper conductors. Except the capacitance C_{34} the variation of the other capacitors is dominated by shielding effects. C_{B1} and C_{21} decrease slightly with increasing distance d_{34} owing to the fact that in this case the upper conductors have a more effective shielding on the lower conductors. On the other hand with increasing distance d_{34}, conductor 4 causes less shielding of field lines originating from conductor 3. Therefore C_{31} increases with in creasing d_{34}. Although the lower conductors and the substrate have a shielding effect on C_{34}, this capacitance still varies with d_{34}^{-1}.

Finally in fig. 8 the insulator thickness d_{13} between the upper and lower conductors has been varied. Owing to a reduction of shielding effects the interline capacitances between conductors at the same level increase considerably with an increase of d_{13}. Naturally the capacitances between conductors at different levels decrease at the same time, but similar to the case of fig. 6 this decrease is only a sublinear function of the thickness.

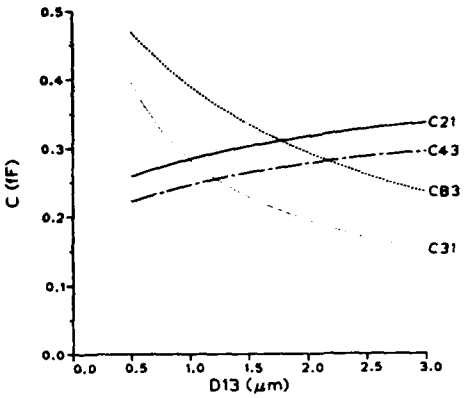


Figure 8. Capacitances vs. insulator thickness ($d_{12} = d_{34} = 2$ μm, $d_{1B} = 1$ μm).

Conclusions

When comparing the results to the limited number of analytic expressions (case 1 and 3), already strong deviations are observed. All trends or effects observed in the figures can be explained from a physical point of view. A change of interconnect dimensions or mutual distances has much larger effect than the thickness of dielectric layers. For 1 μm wide conductors 2-D capacitance values may exceed the 1-D value by a factor of 5, and in 3-D cases by a factor of 9. Estimating 2-D capacitance values by reducing surface charge integrals to line integrals or 3-D capacitance values as a summation of 2-D cases is too inaccurate.

References

[1] M. Elmasry, IEEE **EDL-3**, 1 (Jan. 1982).
[2] T. Sakurai et al., IEEE **ED-30**, 2 (Febr. 1983).
[3] E.W. Greeneich, IEEE **ED-30**, 12 (Dec. 1983).
[4] D.J. Coe et al., Philips proprietary 3-D device simulator.
[5] R. Shrivastava, IEEE **EDL-6**, 3 (March 1985).
[6] P.E. Cottrell et al., IBM J. **R/D-29**, 3 (May 1985).

Session P2.1

Posters

Tuesday, September 15, 1987

IMPLANTATION and DIFFUSION MODELLING of BORON in SILICON

An De Keersgieter, Luc Dupas, Kristin De Meyer*
IMEC, Kapeldreef 75, B-3030 Leuven, Heverlee, Belgium

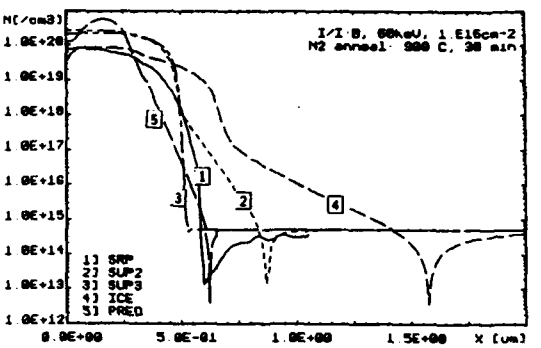
Abstract

The accurate simulation of implanted and diffused impurity profiles in silicon is extremely important when developing VLSI processes. In this work simulations with different process simulators and the corresponding experimental results for implantation and diffusion in N_2 ambient of boron in silicon are compared. Our study reports a remarkable dose dependence of the shape of the experimental profiles. Strategies have been developed to increase the simulation accuracy.

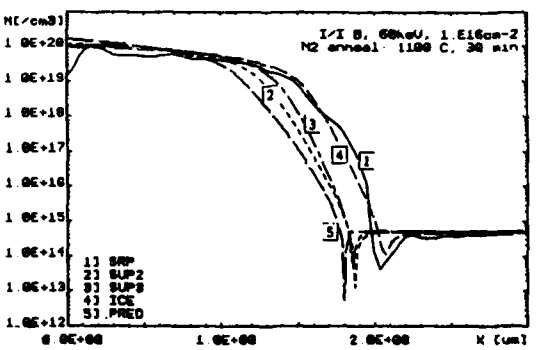
1 INTRODUCTION

The simulators used in this study are: SUPREMII.5[1], SUPREM3[2], ICECREM[3] and PREDICT[4]. The substrates used were 9-13 Ωcm phosphorus-doped n-type (100) silicon wafers. A 400 $\AA$ thick thermal oxide was grown in dry oxygen at 1000°C. Then 60 keV, $^{11}B^+$ implants with doses ranging from 10^{12} cm^{-2} to $10^{17}cm^{-2}$ were performed with wafers tilted 7° to suppress channeling. The annealing temperatures and times were : 900, 1000, 1100 °C and 30, 300, 1200 min.

Comparing the experimental implantation/diffusion profiles and simulations (using default model parameters) no satisfactory agreement is achieved. It is found that for long, high temperature anneal steps simulations and experiments coincide far much better than for low temperature, short time annealing conditions (Figure 1). In the latter case the actual shape of the implantation (starting) profile has a pronounced influence on the final diffusion profile. This is illustated in Figure 2 for profiles simulated with ICECREM. For a 30 min anneal at 1000°C we still see a tail in the profile which is originating from a strong implantation channeling model. After 300 min of annealing the shape of the profile has completely changed. So, the exact modelling of the implantation profile is not only necessary for improving the accuracy of the final diffused profiles, but it is also very important for a better understanding and comparison of the diffusion mechanisms . Also modelling of rapid thermal annealing has a need for precise implantation start profiles.



a)



b)

FIGURE 1 :

Comparison between simulations and experiments
a) for low temperature : 900°C, 30 min
b) for high temperature : 1100°C, 30 min

*Professor of the Katholieke Universiteit, Leuven, Research Associate of the Belgian National Found for Scientific Research.

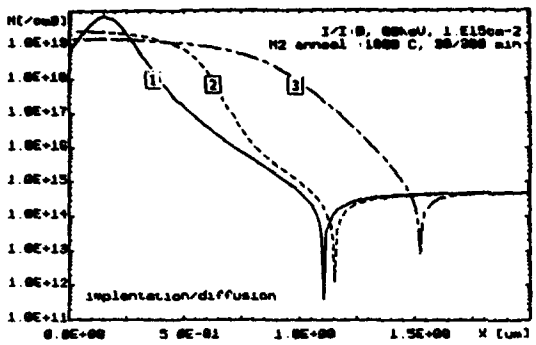


FIGURE 2 :

ICECREM default profiles :
[1] as implanted profile : B, 60keV, 10^{15} cm^{-2} (400 Å oxide)
[2] same profile after N_2anneal, 1000°C, 30 min
[3] same profile after N_2anneal, 1000°C, 300 min

2 DEFAULT IMPLANTATION MODELLING

Most process simulators use analytical formulas for modelling the implanted profiles. Here different models and also different default parameter values are found when comparing the available programs. Most frequently used are the Gauss model (non-channeled profiles), or Gauss-related, PearsonIV and PearsonIV-related (profiles with channeling tail) models. The model chosen depends on the specified impurity, on the type of target, on the thickness of the implant oxide and also on the simulator used (Table 1).

SUPREMII.5	Gauss		PearsonIV + exp.tail
ICECREM	Gauss		PearsonIV
SUPREM3	Gauss	2-Gaussian	PearsonIV
PREDICT			Gauss + exp. tail

Table 1: Different process simulators and their models for implantation of boron in crystalline silicon

The Gauss model

$$C(x) = \frac{Dose}{\sqrt{2\Pi}\sigma} exp\left(-\frac{(x-R_p)^2}{2\sigma^2}\right)$$
$$= Dose \times f(x) \quad , f(x): frequency function$$

gives a symmetrical profile which can only be used for non-channeled implantation profiles (e.g. in amorphous silicon). R_p is the range of the profile and σ is the standard deviation.

In crystalline silicon channeling occurs and the profile is no longer symmetrical. Then at least one has to account for the skewness γ by using a frequency function which consists of two half Gaussian distributions. If also the peakedness β is needed to describe the asymmetry, the PearsonIV distribution is used. This distribution is given by [5]:

$$C(x) = Dose \times f(x)$$
$$and$$
$$f(x) = K[-(b_0 + b_1 y + b_2 y^2)]^{-1/2b_2} \times exp\left(-\frac{b_1/b_2 + 2a}{\sqrt{4b_0b_2 - b^2}} \times arctan\frac{2b_2 y + b_1}{\sqrt{4b_0b_2 - b^2}}\right)$$

a, b_0, b_1, b_2 are functions of σ, γ, β

In this study we concentrate on boron implantations in crystalline silicon, unannealed or annealed in N_2 atmosphere. As all experimental as-implanted profiles (measured by SIMS) show a certain amount of channeling (even with a 400 Å thick implant oxide and 7° tilt) PearsonIV, modified PearsonIV and modified Gaussian models are used (column 4 of Table 1). Using the default implantion models and parameters none of the process simulators gives a satisfactory agreement with the experimental as-implanted profiles over a wide range of doses (Figure 3). SUPREMII.5 and ICECREM show the same tendencies. In both cases the tail extends beyond the extrapolated tail from the SIMS measurement for doses of 10^{15} cm^{-2} and higher. For 10^{13} and 10^{14} cm^{-2} simulations and experiments agree considerably well and for the lowest dose (10^{12} cm^{-2}) simulations seem to underestimate channeling. PREDICT is the only simulator which accounts for a dose dependency for channeling, but the modelling is done in a empirical and discontinuous way. SUPREM3(version 3C) does not model channeling at all. Also remarkable is that literature about SUPREM3 always mentions a PearsonIV model while the default implantation parameters satisfy the conditions for a Pearson I distribution.

Our study thus revealed a remarkable dose dependence of the shape of the experimental profiles (especially in the channeling tail region) which is completely neglected by the available simulators (except to some extend for PREDICT). So it is obvious that the default implantation parameters cannot be used when simulating the actual implanted profiles and as such will also provide a non-optimal starting profile for the study of subsequent diffusion steps.

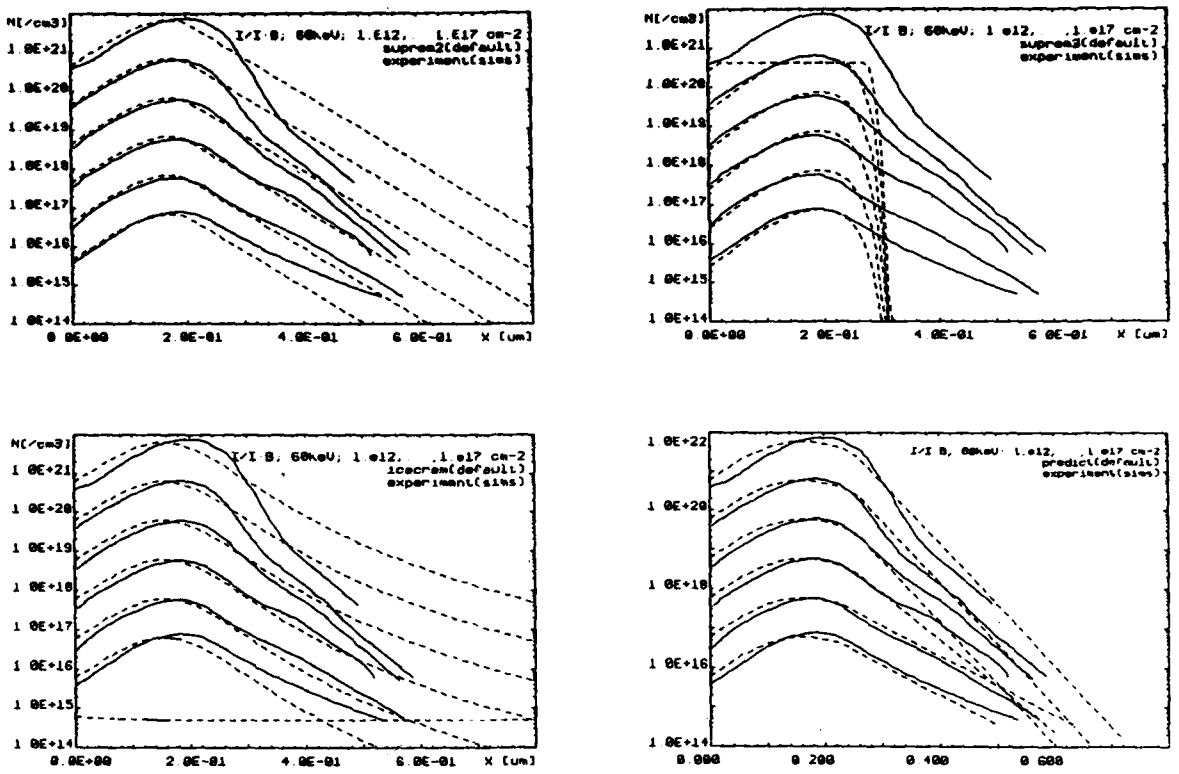


FIGURE 3 :

Comparison between different process simulations and experimental SIMS implantation profiles for I/I: B, 60keV, 10^{12}, 10^{13},...,10^{17} cm^{-2} through 400 $\AA$ oxide. The solid lines represent SIMS-data. The origin of the x-axis is at the Si/SiO_2 interface.

3 NON DEFAULT IMPLANTATION PROFILES

In our work we used the PearsonIV model (used in ICECREM) for extracting accurate model parameters from the experimental SIMS profiles using the SIMPAR[5] package. This parameter extraction program has originally been developed for determining the parameters used in the I-V relations of transistors, but it is also possible to built in any user defined model.

Very good PearsonIV fittings for the experimental implantation results are obtained (Figure 4). The strong channeling effect at higher doses has disappeared, but still some tailing depending on the dose, can be observed. This fitting procedures leaves all distribution parameters ($R_p, \sigma, \gamma, \beta$) free. This is of course not very practical for process simulation, but the resulting sets of parameters do exhibit a dose dependence which can be quite easily described. To model this dependence in a phys-

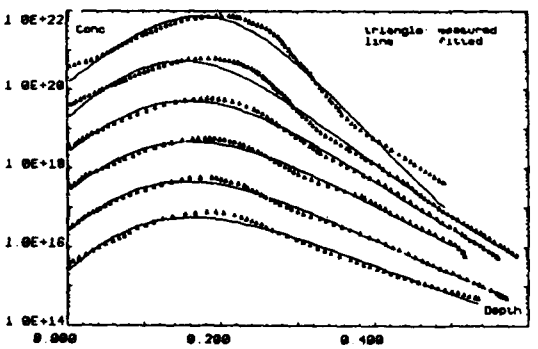


FIGURE 4 :

SIMPAR (PearsonIV) fitting results for implantation

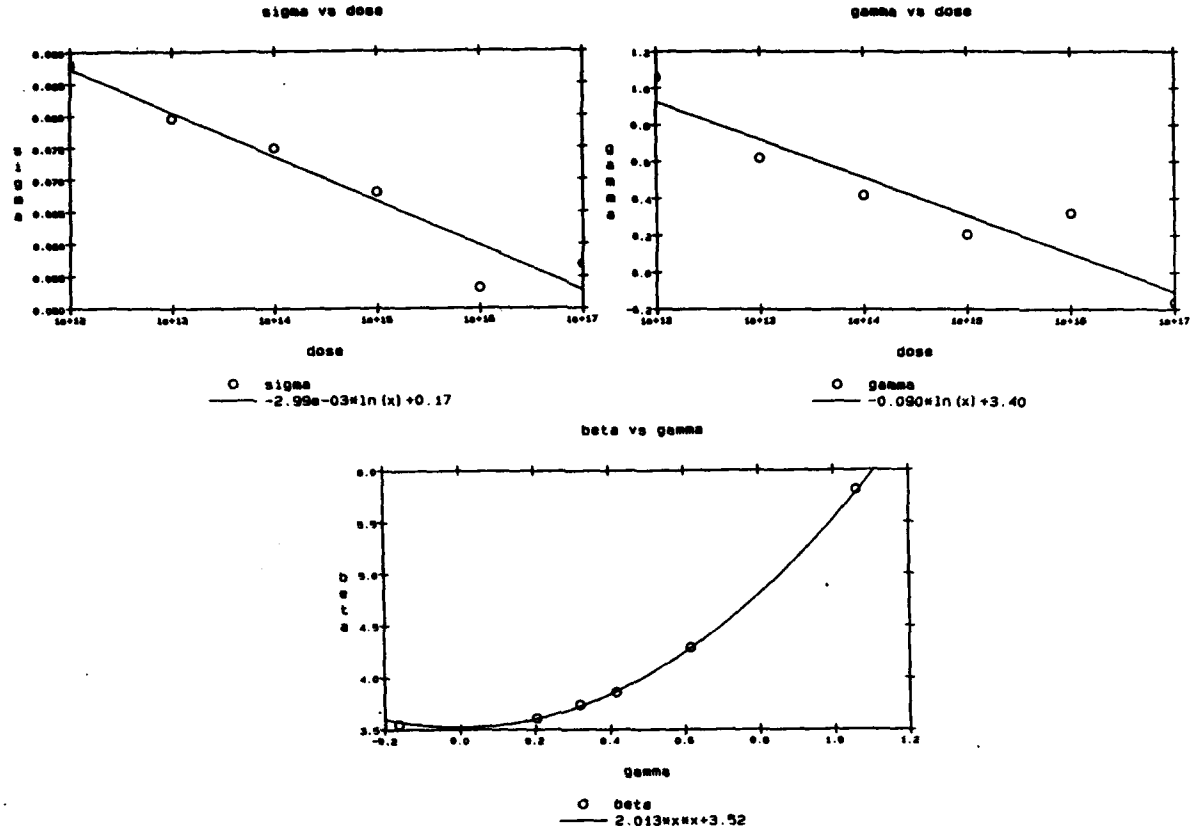


FIGURE 5 :

Dose dependent implantation parameters : σ vs dose, γ vs dose, β vs dose
The circles represent extracted values.

ical and practical way, a second fitting procedure is performed where R_p is kept fixed at a mean value corresponding to the implantation energy. This means we consider R_p independent of the implantation dose. This approach is generally accepted. On the other hand, σ, γ, β are still clearly dose dependent. This explains also the variation of the position of the maximum $= R_p + a$, with a a function of γ and β. The results for a 60keV boron implantation through 400 $\AA$ oxide are given by :

$$\begin{cases} R_p &= cte \\ \sigma &= -2.99 10^{-3} ln(dose) + 0.17 \\ \gamma &= -0.090 ln(dose) + 3.40 \\ \beta &= 2.013\gamma^2 + 3.52 \end{cases}$$

The extracted parameters and the resulting relationsships are shown in Figure 5. All non-constant parameters are decreasing with increasing implantation dose. The skewness γ and the peakedness β are approaching to Gaussian values ($\gamma = 0, \beta = 3$) as the dose increases. This means the higher the dose the less channeling occurs. Indeed for higher doses more damage is created, which reduces the probability for channeling.

4 DIFFUSION PROFILES

Finally starting from the proper implantation profile, different diffusion models can be compared. All diffusion profiles were measured by spreading resistance profiling (SRP). In general we now get a much better agreement with the experimental results, also for the low temperature, short time anneal conditions. Indeed, the strong channeling tail, which could be observed for default simulations and is not at all occuring in the experimental profiles, is eliminated (Figure 6, Figure 7).

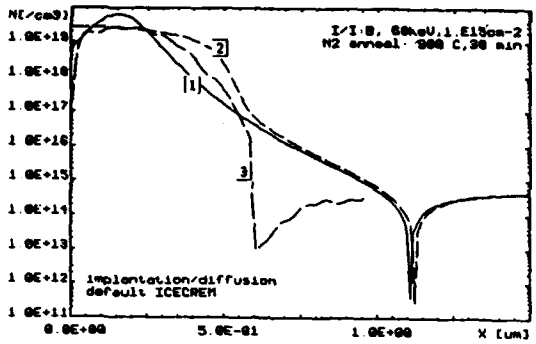


FIGURE 6 :

ICECREM default profiles vs experimental SRP profile :
[1] ICECREM implantation B, 60keV, $10^{14}cm^{-2}$(400 Å oxide)
[2] ICECREM N_2anneal, 900°C, 30 min
[3] experimental annealed profile

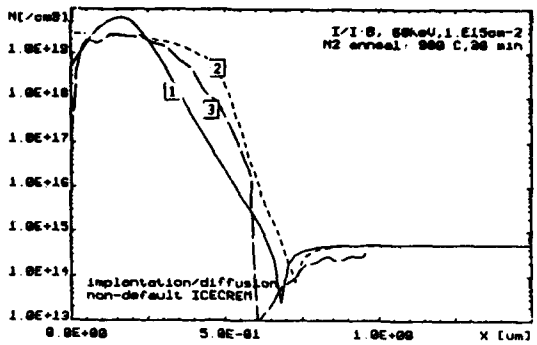


FIGURE 7 :

ICECREM non-default profiles vs experimental SRP profile :
[1] ICECREM implantation B, 60keV, $10^{14}cm^{-2}$(400 Å oxide)
[2] ICECREM N_2anneal, 900°C, 30 min
[3] experimental annealed profile

Simulations prove that due to the bevel angle used for SRP the measured junction is 20% less deep than the real junction in a non-beveled sample. This bevel-effect [7] is also responsable for the kink observed in the SRP-profile.

A more profound basis for the study and improvement of the diffusion model in the different simulators is established. Also this approach provides a more physically realistic strategy for adapting the default model parameters in order to fine tune the simulated results towards the experimental results.

It is made clear that it is very important to look carefully at the existing programs and the implemented models in order to get reliable results. Indeed, apparently good results can be obtained for certain process steps or subsequent process steps, but more critical conditions can cause a failure of the model.

ACKNOWLEDGEMENTS

The authors would like to thank the MAP/ARS group of IMEC for the SIMS and SRP data.

REFERENCES

[1] D.A.Antoniadis, R.W.Dutton, IEEE Journal of Solid State Circuits, vol SC14, No2, April 1979, p412

[2] C.P.Ho, J.D.Plummer, S.E.Hansen, R.W.Dutton, IEEE Transactions On Electron Devices, vol ED-30, No.11, November 1983, p1438

[3] H.Ryssel, K.Haberger, K.Hoffmann, G.Prinke, R.Dümcke, A.Sachs, IEEE Transactions On Electron Devices, vol ED-27, No.8, August 1980, p1484

[4] R.B.Fair, PREDICT, Microelectronics Center of North Carolina (MCNC)

[5] S.Selberherr, Analysis and Simulation of Semiconductor Devices (Springer-Verlag, Wien, 1984)

[6] W.Maes, K.M.De Meyer, L.Dupas, IEEE Transactions on CAD, vol.CAD-5, No.2, April 1986, pp.320

[7] J.Albers, Some Aspects of Spreading Resistance Analysis, in: Gupta,D.C. and Langer,P.H.(eds.), Emerging Semiconductor Technology, ASTM STP 960, (American Society for Testing and Materials, 1986)

SHALLOW JUNCTIONS OF BORON IMPLANTED IN GE^+ PREAMORPHIZED <100> SI WAFERS

A. La Ferla*, S. Cannavò°, G. Ferla°, V. Raineri*, E. Rimini*

* Dipartimento di Fisica, Corso Italia 57-I95129 Catania - Italy.
° SGS - Microelettronica S.p.A., Stradale Primosole 50, I95100 Catania Italy.

p-type shallow junctions in silicon were obtained by preamorphization with Ge^+ ions, 20 KeV B^+ implants at doses in the $5*10^{14}$ - $5*10^{15}/cm^2$ range, and thermal annealing at 850 °C or 950 °C for 1/2 hr. The junction depth was < 0.3 μm in the preamorphized wafers and > 0.3 μm in the B-bare Si implanted wafers. The leakage current density measured at a reverse bias of 10 v was about 100 pA/cm^2 for both procedures. The leakage maps on the 5" wafer gave a density of short circuits of $0.6/cm^2$.

1. INTRODUCTION

The formation of shallow junctions by ion implantation is hampered by channeling effect and by the partial electrical activation in the tail region of the implanted profile in particular for B-type region [1]. *Preamorphization with different ions as* Si^+, Ge^+, Sn^+ of a surface layer deeper than the active range of the p-n structure prevents channeling tails [2-3] and the use of rapid thermal annealing [4] provides a quite complete dopant activation.

The use of Ge^+ ions to preamorphize the layer presents some advantages [5]. The implantation can be performed at room temperature, Ge has an infinite solid solubility in Si, its large covalent radio causes a biaxial compressive strain which can compensate the lattice strain that arises after subsequent high dose implantation and annealing of impurities like B or P. The Ge implantation is superior to the Si^+ implantation in achieving uniform amorphization and a regrown region of high structural perfection results [6].

In the present work we report in detail the electrical characterization of pre-amorphized Si layers with Ge^+ of pre-amorphized Si layers with Ge^+ ions, subsequently implanted with B^+ and then thermally annealed, in view of the relevance of the electrical response of shallow junctions for device applications.

2.EXPERIMENTAL

Silicon wafers, 5" n-type of 1.5 - 4.0 Ω *cm resistivity, of <100> orientation were preamorphized with Ge ions - 150 KeV to a fluence of $5*10^{14}/cm^2$. Some wafers were subsequently implanted with 20 KeV of B^+ to a fluence of $5*10^{14}/cm^2$ (R_p = 670 Å, DR_p = 340 Å). The annealing was performed in a furnace under N_2 flow at temperatures of 850 °C and of 950 °C for 1/2 hr.

The samples were analyzed by 2.0 MeV $^4He^+$ Rutherford backscattering in combination with channeling effect technique. The depth profile of carrier concentration and mobility was obtained by sheet resistence and Hall measurements in layer-by layer removal technique by anodic oxidation. The electrical behaviour was determined by the forward and reverse I-V characteristic and by the leakage maps.

3. EXPERIMENTAL RESULTS

The thermal regrowth of preamorphized Ge^+ layers and subsequently implanted with 20 KeV B^+ was investigated in the 500-600 °C temperature range. The initial amorphous layer amounts to 160 nm. The epitaxial regrowth rate of the amorphous layer is the same for Ge^+ and self-ion Si^+ implantation. The presence of B dopants at concentration of $5*10^{19}/cm^3$ enhances the rate of about a factor ten, in agreement with the data obtained [7] in self-ion implanted layers. Our data support the use of Ge instead of Si, in inducing an amorphous layer. After regrowth Ge atoms occupy substitutional lattice sites.

The sheet resistance of samples implanted with $5*10^{14}/cm^2$ B in either pre-amorphized or virgin <100> Si n-type substrates, and thermally annealed at different temperatures for 30' is shown in Fig. 1.

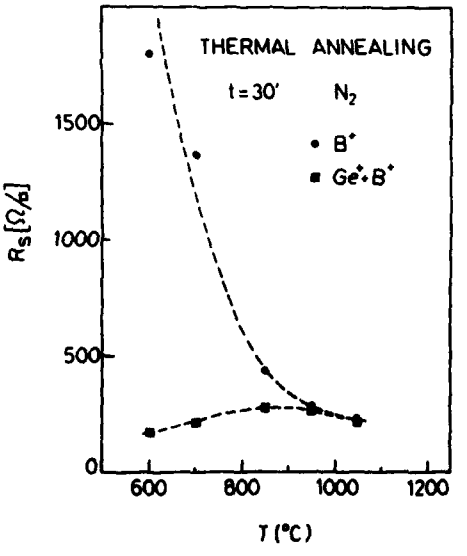


Fig. 1 - Sheet resistance versus annealing temperatures for $5*10^{14}/cm^2$ - 20 KeV B implanted into Ge preamorphized (■) and bare Si (●) respectively.

Sheet resistence is higher in B implanted samples than in B implanted

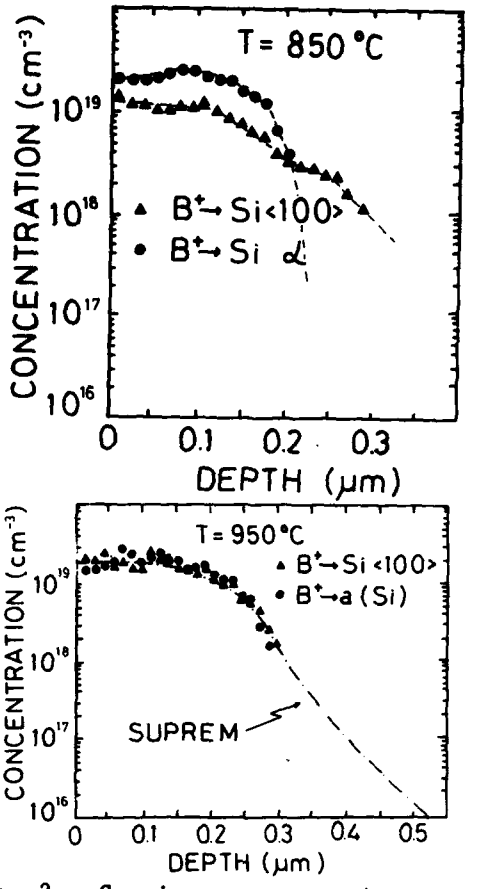


Fig. 2 - Carrier concentration profile of $5*10^{14}/cm^2$ - 20 KeV B^+ implanted into preamorphized (●) and bare (■) Si wafers after thermal annealing for 30' at 850 °C (a) and 950°C (b)

into preamorphized samples up to 900 °C annealing temperature. This behaviour is associated to the regrowth of α-layers and then to the electrical activation of the embedded B dopant at temperatures well below those required to anneal out point and extended defects present in the B implanted samples [8]. The sheet resistance of the layers amounts to 250 Ω/□ in agreement with SUPREM simulation.
The doping and mobility profiles were determined by etching off thin layers of Si one after the other and by measuring the sheet resistance and the Hall coefficient. The thin layers were

removed by anodic oxidation and the Van der Paw geometry was adopted for the electrical measurements. The carrier concentration profiles obtained by this procedure, are shown, in Fig. 2a and 2b for $5*10^{14}/cm^2$ - 20 KeV B^+ implanted into bare Si and into preamorphized <100> Si substrates after annealing at 850 °C and 950 °C respectively. After annealing at 850 °C the activated B^+ dopant fluence amounts to $4.4*10^{14}/cm^2$ in the preamorphized Si and to $6.9*10^{13}/cm^2$ in the bare Si samples in agreement with the resistivity data reported in Fig. 1. The mobility is of 60 $cm^2/v*s$ and 80 $cm^2/v*s$ respectively. After annealing at 950 °C the measured profiles nearly coincide in the investigated range. The activated dose amounts to $4.4*10^{14}/cm^2$ in the preamorphized and $4.3*10^{14}/cm^2$ in the bare Si respectively. The mobility is 60 $cm^2/v*s$ for both. It must be pointed out that the minimum carrier concentration detectable with accuracy by this method is of about $5*10^{17}/cm^2$, i.e. at least two orders of magnitude higher than the substrate doping. The location of the junctions is then determined by the shape of the concentration profiles at values lower than $10^{17}/cm^3$.

As an example the B^+ profile calculated by SUPREM is reported, as full line, in Fig.2b. The tail is approximated in the simulation by an exponential decay, whose slope depends on several parameters, as orientation of the wafer with respect to the beam, annealing temperature, etc. The first part of the simulated profile accounts quite well for the experimental data; but the distribution in the tail cannot be inferred by these measurements, and the location of the junction can be only guessed.

A staining technique has been adopted to measure the junction depth and the following results were obtained after 850 °C and 950 °C annealings: X_j(850°C) = 0.32 μm and 0.2 μm in the bare and in the preamorphized Si, X_j(950°C) = 0.35 - 0.4 μm and 0.28 - 0.3 μm in the bare and in the preamorphized Si. The accuracy of the method is about ± 0.05 μm. The 950 °C seems a good choice for the annealing temperature because at 850 °C a large amount of defects is still present and at higher temperature as 1000 °C for few seconds, a considerable diffusion of the dopant occurs.

The electrical behavior of the formed p-n junctions was investigated by measuring the forward and the reverse I-V characteristics and by the leakage maps in diodes of 0.22 cm^2 area. The reverse cheracteristics are shown in Fig. 3 for the adopted thermal procedures. No pregettering was adopted for the processed diodes. The annealing at 950 °C results in a quite reasonable electrical behavior. The leakage current density measured at a reverse voltage of 10 v was about 100 pA/cm^2 for both procedures (implantation in bare Si or in preamorphized Si). This low value is very close to the theoretical one and indicates that defects and generation centres in the depleted p-n junction are practically absent. The forward I-V characteristcs, not shown, are fitted by an ideality factor $m \sim 1.02$ in the exponential factor.

The leakage current maps were also measured in the processed 5" wafer at a reverse bias of 15 v. A typical map is reported in Fig. 4, together with the histogram of the distribution. The density of short circuits amounts to $0.6/cm^2$. The wafer was processed in such a way that a central region of 2" diameter was preamorphized with Ge and then all the wafer was implanted with 20 KeV B^+. A direct comparison is then possible between preamorphized and as implanted diodes.

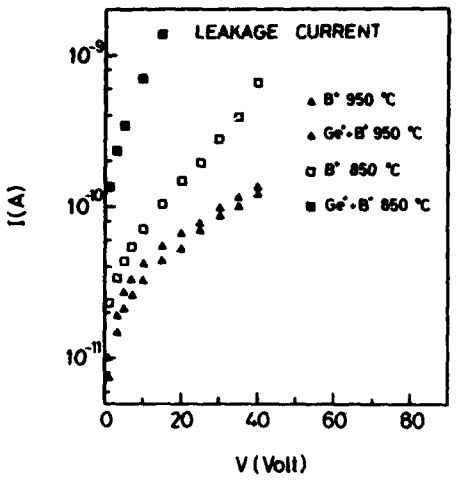


Fig. 3 - Reverse I-V characteristics of 0.22 cm^2 diodes implanted with 20 KeV B^+ - $5*10^{14}/cm^2$.

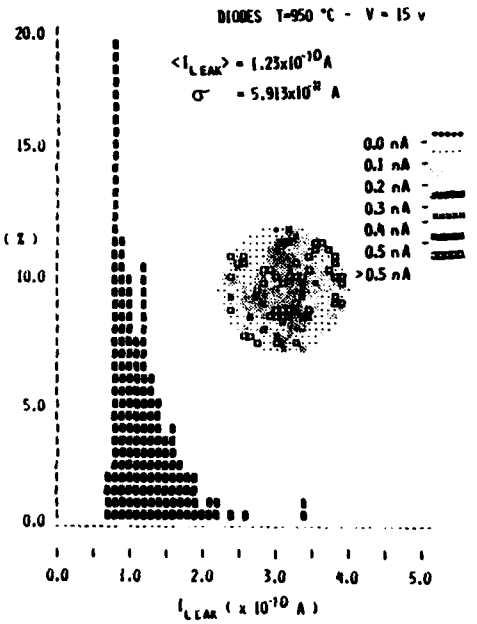


Fig. 4 - Histogram and map of leakage current at - 15 v of diodes annealed at 950°C - 30'; the two inches diameter central region was preamorphized with Ge^+ implantation.

The leakage map and the distribution indicate that within the experimental accuracy no significant difference exists in the electrical behavior of the diodes processed in the two ways. This indicates that the preamorphization with Ge doesn't deteriorate with the residual defects the electrical response.

ACKNOWLEDGEMENTS

The authors wish to tank Dr. M. L. Polignano and Dr. G. F. Cerofolini of the SGS-Microelettronica S.p.A. - Agrate (MI), for an helpful discussion of the electrical results.

Work supported in part by " Programma Nazionale per la Microelettronica ".

REFERENCES

[1] C. Hill, Nucl. Instr. Meth. Phys. Res. B19/20, 345 (1987).

[2] C. Carter, W. Maszara, D. K. Sadana, G. A. Rozgonyi, J. Liu and J. Wortman, Appl. Phys. Lett. 44, 459 (1984).

[3] A. C. Ajmera and G. A. Rozgonyi, Appl. Phys. Lett. 49, 1269 (1986).

[4] G. K. Caller and T. E. Seidel, Appl. Solid State Science-Suppl. 2c editor D. Kahng, Academic Press N. Y. 1985 p. 2.

[5] D. K. Sadana, E. Myers, J. Liu, T. Finstad and G. A. Rozgonyi, Mat. Res. Soc. Symp. Proc. 23, 303 (1984).

[6] S. Prussin, E. R. Weber, K. S. Jones, Nucl. Instr. Meth. Phys. Res. B21, 496-502 (1987).

[7] J. S. Williams in "Surface Modification and Alloying" edited by J. M. Poate, G. Foti and D. C. Jacobson, Plenum Press, N. Y. 133 (1983).

[8] W. K. Hofker, Philips Research Reports, Suppl. N. 8 (1975).

THE EFFECT OF HIGH PRESSURE STEAM OXIDATION ON PHOSPHORUS DIFFUSION IN SILLICON

WU BAI LU

Graduate School of Academia Sinica, China

ZHANG AI ZHEN

Beijing Institute of Semiconductor Devices, China

LI SHY LIN

National Institute of Metrology, China

XUE SHI YING

Academia Sinica, China

The effect of high pressure steam oxidation (7.5-10.7 atm.) on phosphorus extrinsic diffusion in (111)--and (100)--silicon at 700°-970°C has been examined by spreading resistance probes and ellipsometery. It has been found that the OED and ORD appear at the higher and the lower temperature, respectively. The OED-ORD transition point is at about 880°C for 40' in 7.5 atm. for (111)-silicon. The difference in effective diffusion coefficients between oxidation and non-oxidation regions (D) is proportional to $(x_0/t)^n$, the power figure n is related to the oxidation conditions. These results can be explained satisfactorily by considerations which take into account oxidation rate and concentration effect on phosphorus diffusion in silicon.

With the development of VLSI, the smaller device dimensions are needed. Therefore the fine control of impurity profiles is more important. The high pressure oxidation is an excellent oxidation process for VLSI. A precise understanding of the effect of this oxidation on impurity redistribution is a requirement for device processes and has important significance for further solving the physical mechanism of oxidation enhanced diffusion (OED) and oxidation retarted diffusion (ORD).

This paper reports that several new phenomena found under the high pressure oxidation are different from those in atmospherical oxidation.

1. EXPERIMENTAL CONDITIONS

Original silicon wafers were 7-15Ω·cm, p type (100)- and (111)- orientated single crystal silicon. The phosphorus diffusion were performed with sp_1 planar solid source at 1140°C for 8'. After that the surface phase layers were removed with the dilute HF solution. Then the Si_3N_4 films of 1000 Å thickness were deposited on the silicon surface and subsequently were photolithographically patterned into 1 m.m. wide parallel stripes so that both the nonoxidizing and oxidizing regions could be on the same wafer. The thermal oxidation was carried out in a high pressure steam oxidation system which was automatically controlled by a microprocessor. Before and after oxidation, the spreading resistance profiles (R-X) were measured by ASR-100B spreading resistance probes. The oxide thickness (x_0) were measured by a TP-77 ellipsometer.

2. EXPERIMENTAL RESULTS AND DISCUSSIONS

The oxidation curves of (111)-Si in 7.5

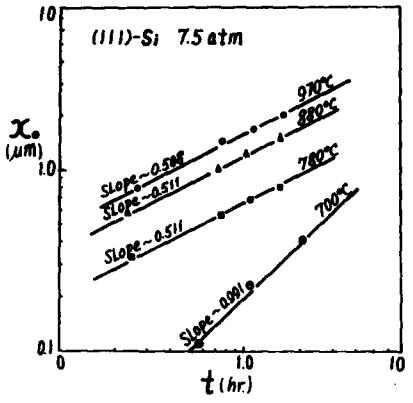


FIGURE 1
Oxidation curves of silicon in high pressure steam at 700^0-970^0C

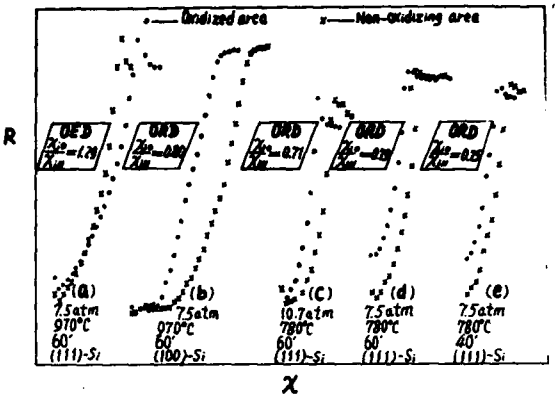


FIGURE 2
Matched spreading resistance profiles of phosphorus in silicon

atm. steam at 700^0-970^0C are given in Fig.1. It can be seen that the oxidation rate (dx_0/dt) is very fast at 970^0, 880^0, 780^0C, it has run in the parabolic oxidation regime. While at 700^0C, it is in the linear oxidation regime. It is implied that the difference in impurity diffusion in silicon under high pressure and atmospherical oxidation is not due to the difference in the oxidation equation, but is correlative with the oxidant supply.

Several typical spreading resistance profiles for different experimental conditions are shown in Fig.2. From 2a and 2b, it can be observed that the impurity distribution and the ratio of the junction deepth, x_{jo}/x_{jn}, are related to the crystalline orientation of samples. Within the pressure and the temperature range used in this experiment, all of the phosphorus diffusion in (100)-Si are retarded (ORD), and its diffusion coeffecients are all bigger than those of (111)-Si at the same condition. By comparing 2c with 2d, it is obvious that the effect of oxidation on diffusion is correlative with pressure (p) of oxidant.

A plot of the logarithm of diffusion coefficient versus the reciprocal of absolute temperature is shown in Fig.3 for phosphorus in (111)-silicon in 7.5 atm. Here, the diffusion coefficients D_O, D_N were calculated from the measured junction depths x_j and sheet resistances in both the oxidation and the nonoxidation region according to Gaussian distribution. By doing so, the segragation effect of the moving boundary in the oxidation region is ignored. The validity has been discussed by some authors [1, 2, 3]. It can be noted from this figure that $D_O > D_N$ at the higher temperature (970^0C), that is OED. While $D_O=D_N$ at 880^0C for 40', it is the transition point of OED-ORD. When the oxidation time t is over 40' or the temperature is lower than 880^0C, then $D_O<D_N$, that is ORD. Another regular phenomenon is that D_O decreases as increasing the oxidation time t at the same temperature, and D_N also decreases except for 970^0C so that the enhancement decreases while the retard increases.

A plot of ΔD versus time-mean effective oxidation rate x_0/t is shown in Fig.4. A result of normal dry oxygen oxidation given by Y. Ishikawa [9] is also plotted in the figure. From Fig.4, ΔD and x_0/t have a power relationship except for near the transition point.

$$|\Delta D| = A(x_0/t)^n$$

within the parabolic oxidation region indicated in Fig.1.

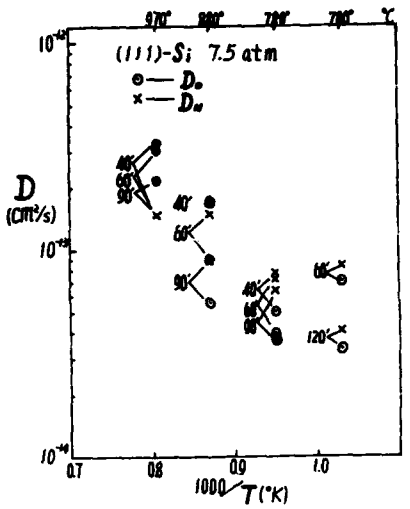


FIGURE 3
Diffusion coefficient of phosphorus in silicon versus reciprocal temperature

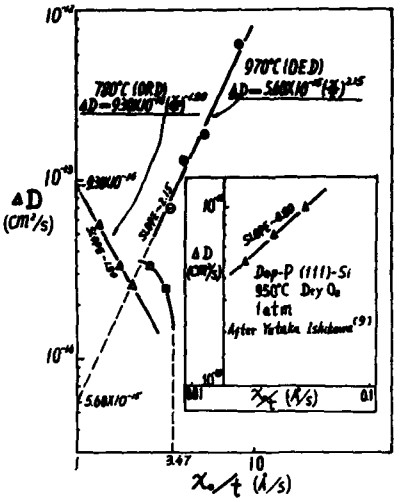


FIGURE 4
Oxide growth rate dependence of D of phosphorus in silicon

$$x_0/t = 2(dx_0/dt)$$

Thus, we have

$$|\Delta D| \propto (dx_0/dt)^n$$

Where, the n is different from the predictive value (n=1) of S.M.Hu model [7] and the results given at normal oxidation by other authors [9, 10]. The n can be positive or negative. It is depend on the temperature, pressure, oxidant composition and crystalline orientation of samples. Particularly near the OED-ORD transition point, the n approaches infinity as a limit. It means some criticality and is well worth notice.

Another interesting phenomenon is the relationship of $D_O/D_N - \frac{1}{T}$ plotted in Fig.5. The right half is our result in 7.5 atm., but the left half was obtained in dry O_2 by Francis and Dobson [11]. When the temperature is increasing from 700°C to near 1300°C, it undergoes two transitions from ORD to OED and again from OED to ORD. By comparing it with the relationship of the length of oxidation induced stacking faults (OSF) Vs 1/T of Fig.6, it can be found immediately that OSF also undergoes two transitions from the shrinkage to the growth, again to the shrinkage in the similar temperature range. These results verify once again that OED (ORD) and OSF are two behaviours coming from a same physical process - the injection of non-equilibrium point defect, and the phosphorus diffusion in silicon is mainly via interstitialcy mechanism.

The measured effective diffusion coefficient (D) may be considered to be a superposition of three physical effects: the Fick effect Di, the concentration effect ΔDc and the surface oxidation effect ΔD_O, that is

$$D = Di + \Delta Dc + \Delta D_O$$

In the nonoxidation region, $\Delta D_O = 0$. Also $\Delta Dc(970^0C) = 0$, since the surface impurity concentration ($1.3 \times 10^{19} cm^{-3}$) is lower slightly than the intrinsic carrier concentration in silicon at 970°C. Thus, $D_N(970^0C) = Di$ and has no change with t. While below 970°C, $\Delta Dc \neq 0$ and change with t, and so do the D_N. In the oxidation region, there are both ΔDc and ΔD_O. Strictly speaking, ΔDc and ΔD_O effect each other [4]. But, as a first approximation, we can assume [5, 6] $\Delta D_c(O) \simeq \Delta D_c(N)$, then

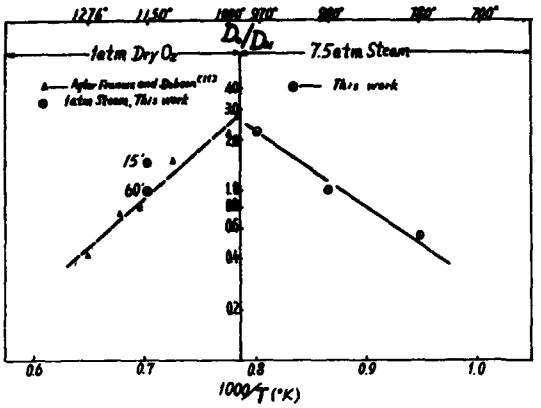


FIGURE 5
Ratio of the diffusion coefficients (D_O/D_N) versus reviprocal temperature

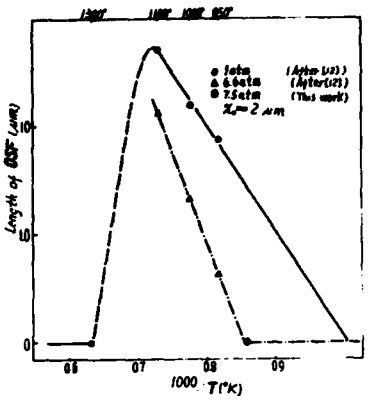


FIGURE 6
Length of OSF versus reciprocal temperature

$\Delta D=D_O-D_N=\Delta D_O$, that is, the measured ΔD can be approximately considered to be the result from oxidation solely.

The ΔD_O, as has been recognized [7, 8], come from non-equilibrium point defect injection into silicon caused by oxidation reaction at SiO_2-Si interface. In turn, the point defect injection would be dependent upon the relative rate between the diffusion transporting of oxidant at SiO_2-Si interface and the reaction consuming of it at the interface. This suggests that the lack and the excess of oxidant at the interface result in interstitail and vacancy injection, and so for OED and ORD of phosphorus. The pressurizing is in favour of oxidant transporting, the decrease of temperature and the increase of time result in retard of interface reaction, hence ORD is favoured. In a similar way, the difference in oxidant composition and crystalline orientation also has effects on OED (ORD).

In summary, the result of high pressure oxidation shows that the phosporus diffusion can be enhanced or retarded, it is dependent upon the oxidation condition.

REFERENCES

[1] Wu Bai-Lu et.al., ICSICT'86, Beijing, p.70

[2] T.Kato et.al., Jpn, J.Appl. Phys. 3(7), 377 (1964).

[3] Y.Nabeta et.al., J.Electrochem. Soc.,123, 1416 (1976).

[4] B.E.Deal et.al., J.Electrochem. Soc.,112, (4), 430 (1965).

[5] C.P.Ho et.al., J.Electrochem. Soc., 126(9), 1516 (1979).

[6] Wu Bai-Lu et.al., to be published.

[7] S.M.Hu, J.Appl. Phys. 45(4), 1567 (1974).

[8] T.Y.Tan et.al., Appl. Phys. Lett. 40(7), 1. April, 1982.

[9] Yutaka Ishikawa et.al., J.Electrochem. Soc. 129(3), 644, 1982.

[10] K.Taniguchi et.al., J.Electrochem. Soc., 127(10), 2243, 1980.

[11] R.Francis and P.S.Dobson, J.Appl. Phys. 50(1), 280, 1979.

[12] Natsuro Tsubouchi et.al., Jpn.J.Appl. Phys. 17, 223-228 (1978).

SHALLOW JUNCTION FORMATION USING $CoSi_2$ AS A DIFFUSION SOURCE

V. Probst[1], P. Lippens, L. Van den hove, K. Maex, H. Schaber* and R. De Keersmaecker

Interuniversity Microelectronics Center (IMEC v.z.w.), Kapeldreef 75, B-3030 Leuven, Belgium
***SIEMENS AG, Central Research and Development, Otto-Hahn-Ring 6, D-8000 Munchen 83, BRD**

Thin layers of $CoSi_2$ (120 nm) were used as a source for B and As diffusion in order to form shallow steep junctions with high interface concentration. SIMS depth profiling as well as two-dimensional characterisation of the indiffusion demonstrate the power fo this technique over a wide range of temperatures and times. Diodes with a high yield and a very low leakage current density ($\simeq$1 nA/cm^2) prove the reliability of the process.

1. INTRODUCTION

The shrinking of device dimensions in integrated circuits necessitates a significant reduction of junction depths towards the 0.1μm scale. In order to improve device and circuit speed, however, parasitic elements such as series resistance have to be reduced as well. These two requirements cannot be satisfied simultaneously by conventional junction formation processes, nor by dopant diffusion from poly-Si such as applied e.g. in self-aligned bipolar devices [1]. Silicides, on the other hand, are proven to be successful in order to reduce the contact and sheet resistances. The silicidation of extremely shallow diodes, however, requires special attention. When the conventional silicidation of preformed shallow junctions is applied, the diode yield decreases drastically with decreasing junction depth/silicide thickness ratio [2]. Therefore, the diffusion of dopants from the silicide into the silicon is an attractive alternative to this process [3]. Besides the benefit of avoiding direct implantation of dopants into the single-crystal silicon, the proposed method of junction formation is expected to be 'self-adjusting' to the silicide/silicon interface shape which has a certain degree of roughness. Therefore, junction shortage due to inhomogeneous silicidation can be avoided.

2. EXPERIMENTAL

In this work, $CoSi_2$ was used as a source for B and As diffusion into mono-Si. In order to study the fundamental characteristics of this diffusion source, the $CoSi_2$-salicide process, as described in [2], was applied first on unpatterned Si-wafers. After formation of 120 nm $CoSi_2$ at 700°C for 30 s (RTP), the desired dopant was implanted into the silicide with a dose of 5.E15 cm^{-2}. Simulations using the program TRIM85 [4] were used to determine a suitable implantation energy assuring confinement of the implant within the $CoSi_2$-layer (20 keV in the case of B, 50 keV for As). Prior to the diffusion step, the silicide was capped with 200 nm of CVD-SiO_2 in order to prevent dopant loss to the ambient. The diffusion cycle was either performed in a conventional furnace or in a rapid thermal processing system (RTP), in order to test the diffusion behaviour and stability of the $CoSi_2$-source over a wide range of temperatures and times.

In a second experiment shallow silicided n^+-p and p^+-n diodes were fabricated using the diffusion of dopants from $CoSi_2$. Active areas were defined by a conventional LOCOS-technique. The silicidation, implantation and diffusion steps were carried out following identical conditions as for the unpatterned wafers. Then a CVD-oxide was deposited and contact windows were opened. Ti/W was used as a diffusion barrier between the silicide and the Al-metal layer. Finally, a sintering in forming gas at 450°C for 30 min was carried out.

3. RESULTS

Four-point probe measurements were performed between the processing steps to check the change in $CoSi_2$ sheet resistance. Due to implantation damage, the initial sheet resistance

[1]Permanent address: Siemens AG, Central Research and Development (München)

of 1.25 Ω/□ after silicidation [2] increased to 5 Ω/□ in the case of B implantation and 3 Ω/□ for As. The subsequent diffusion steps anneal this implantation damage and cause the complete recovery of the $CoSi_2$ sheet resistance (e.g. 1.1 Ω/□ after 950°C anneal for 30 min), which is in good agreement with [5].

The diffusion of the dopants in the silicide was measured by secondary ion mass spectrometry (SIMS). Figure 1 shows the as-implanted B profile in comparison with the 800°C/30 min diffusion step. Due to the fast diffusion of B in $CoSi_2$ at that temperature a complete equidistribution has taken place in the silicide at a concentration level of about 3.E20 cm^{-3}. At the interface to the SiO_2-layer, boron tends to segregate (SiO_2 was etched off prior to SIMS measurement).

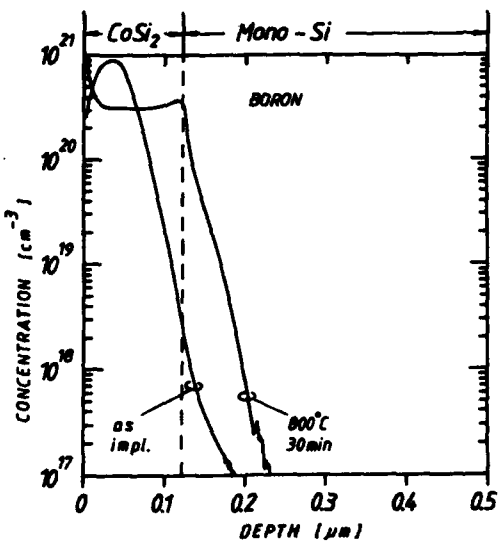


FIG. 1: Boron profiles measured by SIMS: as-implanted and diffused out of $CoSi_2$ at 800°C for 30 min.

In fact, more important for device applications is the indiffusion of the dopants from the silicide into the mono-Si. Especially the temperature-dependent junction depth and interface-concentration are of main interest for scaling and contact-resistance respectively. In order to exclude matrix-effects during SIMS-measurement and to avoid limitations in depth resolution, the $CoSi_2$ was selectively removed from the mono-Si by etching in 25 % HF. Figure 2 shows the indiffusion of boron from $CoSi_2$ for different heat-cycles. Junction depths from 100 nm (for 800°C/120 min) to 500 nm (for 1100°C/60 s) and

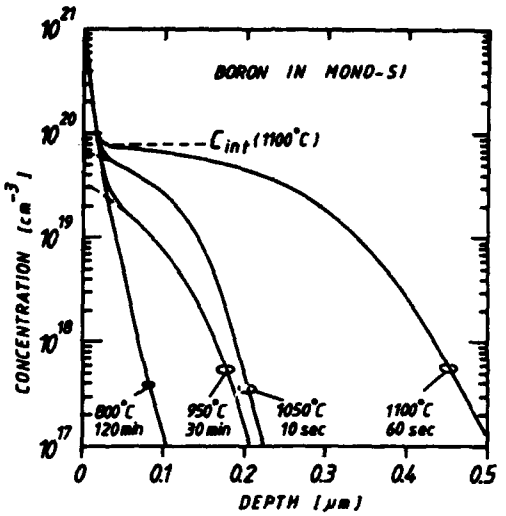


FIG. 2: SIMS-profiles of B-indiffusion from $CoSi_2$ at different heat cycles.

interface-concentrations between 3.E19 cm^{-3} (at 950°C) and 8.E19 cm^{-3} (at 1100°C) show the reliability of $CoSi_2$ as a diffusion source over a wide temperature and time range. A junction depth of about 200 nm can either be achieved by 950°C/30 min furnace anneal or by 1050°C/10 s RTP. As expected, the surface concentration is higher for the RTP sample. The same tendency is shown in fig. 3 for As-indiffusion. However, due

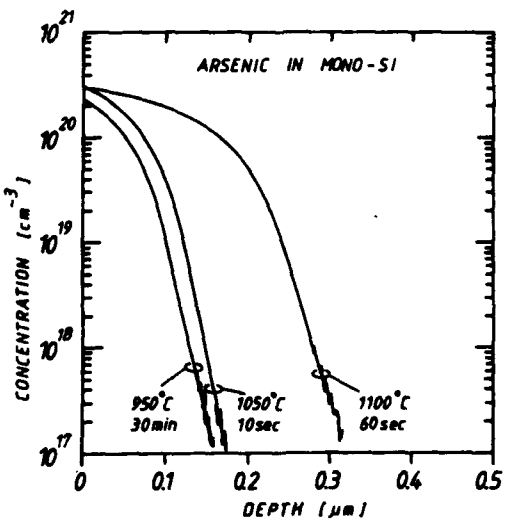


FIG. 3: SIMS-profiles of As-indiffusion from $CoSi_2$ at different heat cycles.

to the lower diffusivity of As in mono-Si [7], the As-junctions are shallower than the B-junctions formed at equivalent conditions. Interface concentrations between 2.E20 cm^{-3} and 3.E20 cm^{-3} were achieved for As at 950°C and 1100°C, respectively. In general, the indiffusion behaviour of B and As from $CoSi_2$ is very similar to the indiffusion from poly-Si [6], [7]. Slight deviations in the interface concentrations could be due to the different segregation behaviour or solid solubility of dopants in $CoSi_2$ which requires additional experiments.

SIMS-depth profiling on shallow junctions (< 50 nm) (e.g. 800°C/120 min As-indiffusion) need special measurement conditions such as low energy primary ion beam, which increases measuring time drastically. For determination of the junction depth only, the method of bevel and staining (BS) is more efficient on these samples. Table 1 gives a summary of the junction depths obtained for B and As for various annealing conditions, measured with SIMS, BS and spreading resistance probe (SRP). Taking into account that these independent methods yield junction depths at different doping levels, a good agreement is obtained between the different results which proves the reliability of the data.

	ARSENIC					BORON				
DIFFUSION STEP	800°C 2 h	900°C 30'	950°C 30'	1050°C 10"	1100°C 60"	800°C 2 h	900°C 30'	950°C 30'	1050°C 10"	1100°C 60"
$X_{j\,SIMS}$ [nm] at 1.E17 cm^{-3}			150	170	310	110	150	200	220	500
$X_{j\,BS}$ [nm]	40	70	140	120	290	110	130	190	190	430
$X_{j\,SRP}$ [nm]			160		330			110		470

TABLE 1: Comparison of junction depths obtained by indiffusion from $CoSi_2$ and measured with SIMS, SRP, bevel and staining (BS).

An important issue is the lateral homogeneity of the junction. Since the mentioned analysis methods give one-dimensional information (averaged over a large sample area) only, SEM was applied on selectively etched cross-sections to visualize the $CoSi_2$-grains and the two-dimensional shape of the diffusion front with high resolution (at a dopant concentration of about 5.E18 cm^{-3}).

Figures 4a and 4b give examples for the case of B-diffusion at 900°C/30 min and 1050°C/30 s respectively. The $CoSi_2$-

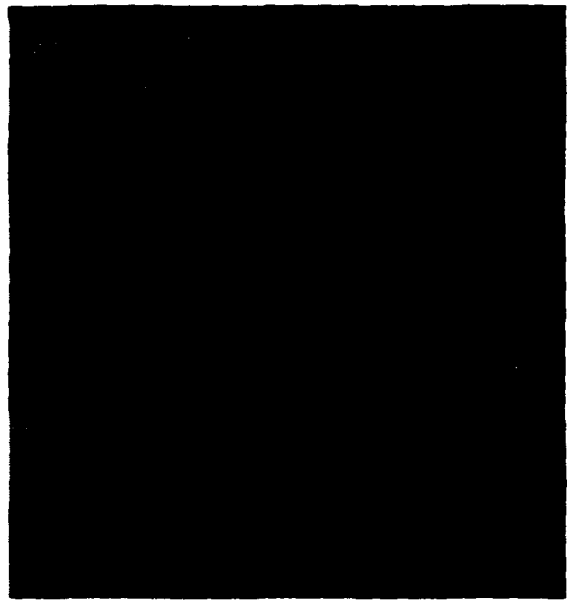


FIG. 4a: SEM-cross-section showing the $CoSi_2$/Si interface and the diffusion front for the boron diffusion at 900°C, 30 min. FIG. 4b: SEM-cross-section for boron indiffusion at 1050°C for 30 s. FIG. 4c: SEM-cross-section showing $CoSi_2$-'globules' surrounded by the diffusion front for As-diffusion at 1100°C, 60 s.

grains, the interface to the mono-Si and the diffusion region are clearly seen. These micrographs reveal that the diffusion front follows the $CoSi_2$/Si interface at a nearly constant distance. The explanation for this self-adjusting mechanism is similar to that for the poly-Si diffusion source in ref. 7. The $CoSi_2$/Si interface itself behaves like a grain boundary with high diffusivity. The dopants are mainly supplied by the vertical grain boundaries, but also from the bulk of the $CoSi_2$ [8]. Reaching the interface, they immediately redistribute in the lateral grain boundaries before slowly diffusing into the mono-Si.

Figures 4a-4b also show that the lateral growth of the $CoSi_2$-grains from $d_G \simeq 0.3\ \mu m$ at 900°C/30 min to $d_G \simeq 0.6\ \mu m$ at 1050°C/30 s is accompanied by a local conglomeration of the silicide. This in consequence causes a temperature-dependent local change in thickness and an increase of the interface-roughness from about 50 nm at 900°C/30 min to about 140 nm at 1100°C/60 s. This is even more pronounced in the case of As-doping, where the $CoSi_2$-layer has balled up locally after a heat cycle of 1100°C/60 s. However, even this extreme case confirms the self-adjusting mechanism of indiffu-

sion by showing the $CoSi_2$-'globules' surrounded by the diffusion front.

Figure 5 gives a typical I-V (forward and reverse) characteristic of diodes formed by a diffusion of B or As at 800°C/120 min in N_2. Table 2 summarizes the average leakage current density for several diffusion conditions (devices with leakage current density higher than 20 nA/cm^2 were considered defective). Even for the highest temperature cycles (1100°C) where the silicide has balled up locally (fig. 4c) leakage currents as low as 1 nA/cm^2 (at 5 V reverse bias) were observed. This also indicates that the generation of trap centers related to Co in the space charge region must be much less pronounced than expected from diffusion length (about 160 μm) and solid solubility (about 1.E15 cm^{-3}) [9] of Co in Si at 1100°C/ 60 s. Due to the non-planar shape of the junction (fig.5) , the breakdown voltage is expected to decrease [10]. However, no difference was seen in comparison with the values obtained from the non-silicided control devices (20 V for As, 30 V for B).

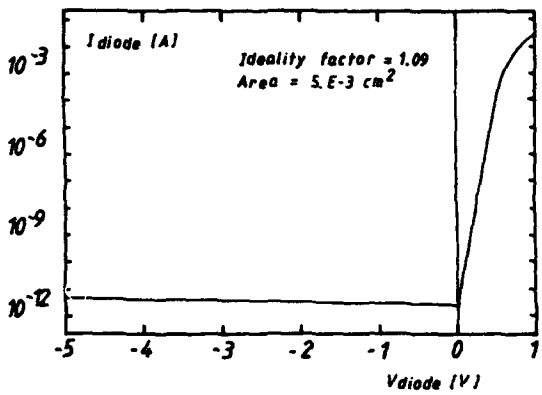


FIG. 5: I-V characteristic of n^+/p-diode: As diffused out of $CoSi_2$ at 800°C for 120 min.

	ARSENIC		BORON	
	mean [nA/cm^2]	% good devices	mean [nA/cm^2]	% good devices
800°C, 120 min	0.9	100	3.3	100
850°C, 60min	1.3	96	3.6	100
950°C, 30min	0.9	100	2.4	96
1100°C, 10 s	3.8	80	1.8	100

TABLE 2: Comparison of leakage current and yield for B and As-diodes formed by indiffusion from $CoSi_2$.

4. CONCLUSIONS

It is demonstrated that $CoSi_2$ is a very useful and flexible diffusion source for arsenic and boron. SIMS-analyses show that shallow as well as deep junctions with high interface concentrations can be obtained by furnace anneal and RTA. Two-dimensional analyses of the diffusion (SEM) reveal an interface-related self-adjustment of the diffusion front which avoids junction shortage due to inhomogeneous silicidation. Diodes formed by B or As indiffusion from $CoSi_2$ show ideal forward and reverse characteristics with very good yield even for extremely shallow junctions of 40 nm (As) and 100 nm (B).

ACKNOWLEDGEMENTS

The authors would like to thank P. Eichinger and W. Vandervorst for SIMS-analyses, K. Wittmaack for useful discussions and also R. De Koninck for device-measurements. P. Lippens is indebted to the Belgian Institute for Scientific Research in Industry and Agriculture (IWONL), whereas L. Van den hove and K. Maex are supported by the Belgian National Fund for Scientific Research (NFWO).

REFERENCES.

1. A. W. Wieder, IEDM Tech. Dig. 1986, p.8
2. L. Van den hove et al., IEEE Trans. El. Dev., ED-34, (March 1987) 554
3. R. Liu et al., presented at the ECS-meeting, May 10-15 1987 (to be published)
4. J. P. Biersack and L. G. Haggmark, Nuclear Instruments and Methods 174, (1980) 257
5. B. M. Ditchek et al., presented at the MRS-meeting, april 1987 (to be published)
6. H. Schaber et al., J. Appl. Phys. 58, (1985) 11
7. V. Probst et al., Semiconductor silicon, (1986) 594
8. P. Gas et al., to be published in J. Appl. Phys.
9. H. Kitagawa et al., Jap. J. Appl. Phys. 21, (1982) 276
10. A. S. Grove, Physics and Technology of Semiconductor Devices, Ed. J. Wiley New York, (1967) 195

2-D EFFECTS DURING ISOLATION PROCESS: EXPERIMENTS AND SIMULATION

A. Seidl

Institut für Festkörpertechnologie
Paul Gerhardt Allee 42
D-8000 München 60

V. Huber

Siemens AG, ZT ZFE FKE 41
Otto-Hahn-Ring 6
D-8000 München 83

SEMIROX bird's beaks were processed for different sets of parameters including variation of the buffer oxide thickness, nitride thickness and temperature. A numerical simulator is used to discuss two-dimensional stress-and diffusion effects.

1. INTRODUCTION

For many standard oxidation processes it can be observed that the oxidation rate is strongly stress-dependent. A qualitative description of this effect has been given in [1,2]. The first effort for a quantitative analysis of this effect was made in [4] where circular etched silicon structures were oxidized and the oxide thickness was measured as a function of the radius. By using rotational symmetry the actual two-dimensional problem was reduced to one dimension and thus could be described by an ordinary differential equation. In [4] the different oxidation rates for concave and convex corners were attributed mainly to a pressure dependence of the viscosity thus allowing no conclusion on the behaviour of oxidant diffusion under stress.

As will be shown in the next section the length-to-width ratio of a bird's beak is a function of the diffusion coefficient/reaction constant ratio. Thus the bird's beak experiment provides better orthogonality between the oxygen diffusion and reaction because it delivers as results not only the oxide thickness but also the length-to-width ratio of the under-diffusion region.

2. 2D DIFFUSION EFFECTS

In [3] it was shown by a simple qualitative model that the length of the bird's beak grows with increasing buffer oxide thickness and diffusion constant and drops with increasing reaction rate. In this case the under-diffusion of the mask via the buffer oxide was considered.

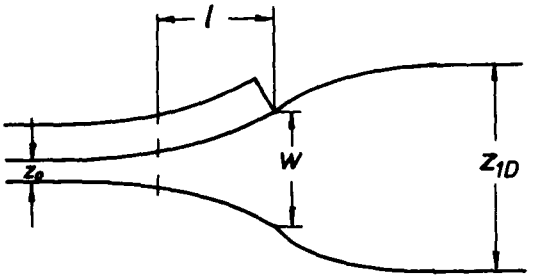


FIG. 1: Typical bird's beak geometry. The start of the under-diffusion-region is defined to be the position where the oxide thickness reaches $z_0+z_{1D}/10$. The width of this region is measured at the mask edge.

Exact numerical values for the dependence of the bird's beak length on D, K, and z_0 can be extracted from numerical simulation only. For this purpose a geometrical definition of the width-to-length (w/l) ratio was defined according to Fig. 1. For the numerical simulation the Deal-Grove model was ex-

tended to two dimensions by using the Finite Element method. Viscous flow was assumed for oxide deformation.

Fig. 4 shows that the dependence of the w/l-ratio on buffer oxide thickness is linear whereas it's dependence on the reaction/diffusion (k/D) ratio exhibits a rather nonlinear behaviour (Fig. 2).

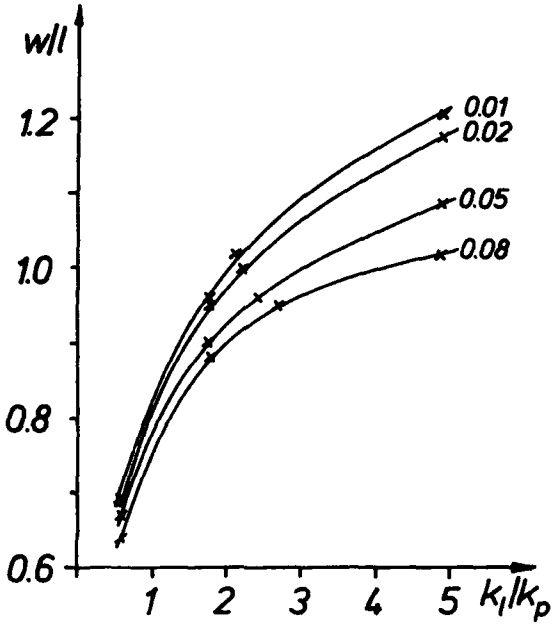


FIG. 2: The w/l ratio of a bird's beak depending on the reaction rate/diffusion coefficient ratio. (The linear and parabolic growth rates are related to these constants by $k_l/k_p = k/2D$)

3. STRESS EFFECTS

The shape of the bird's beak is influenced by mechanical stress. The two mechanisms involved are:

- direct deformation of the oxide by pressure exerted by the nitride mask
- stress-dependent coefficients (Diffusion, reaction, viscosity etc.)

The second set of mechanisms has been observed by various workers but a quantitative model has been given only by Kao [4]. The coefficients of the governing equations are dependent on mechanical stress by a Boltzmann type relationship.

$$(1) \quad k_s = k_o \exp(-\sigma_n V_k/kT$$
$$D = D_o \exp(-pV_D/kT)$$
$$C^* = C^*_o \exp(-pV_c/kT)$$
$$\mu = \mu_o \exp(\alpha(T)p)$$

σ_n denotes the normal stress at the $Si-SiO_2$ interface, p the hydrostatic pressure, C^* the saturation concentration and μ the viscosity of the oxide. The V's are denoted as activation volumes.

4. EXPERIMENTS

In this work <100> lightly n-doped (5-20 Ωcm) silicon wafers were cleaned and oxidized in dry O_2 to form a buffer oxide between 100 and 800 Å. Silicon nitride of a thickness between 200 and 2000 Å was deposited. After chemically etching the nitride a wet oxidation was performed at temperatures between 900 and 1100 °C. The TEM preparation of the cross section was performed by mechanical lapping and successive ion etching.

5. VARIATION OF PAD OXIDE THICKNESS

Samples were fabricated with various pad oxide thicknesses at 1000°C oxidation temperature with a 1200 Å nitride mask. Qualitatively the profiles behave as expected and show reduced under-diffusion for lower pad-oxide thicknesses. However a quantitative comparison between measurement and stress-free simulation shows that the measured bird's beaks have the tendency to be longer



a



b

FIG. 3: TEM micrographs of bird's beaks grown with 800 Å (a) and 100 Å (b) buffer oxide together with stress-free simulation.

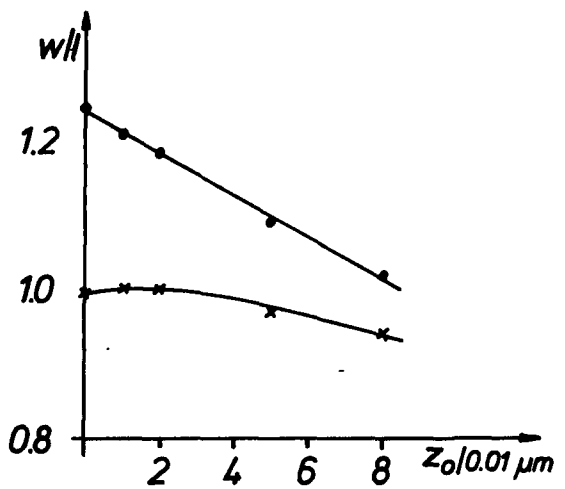


FIG. 4: w/l ratio versus buffer oxide thickness: x=measured, o=stress-free simulation

and flatter than the simulated profiles. The agreement is quite good for a large pad oxide thickness (Fig. 3a). However for small pad oxide thicknesses the difference is clearly visible (Fig. 3b). In Fig. 4 the w/l ratios of measured and simulated bird's beaks are plotted against the buffer oxide thickness. The comparison with the stress-free simulation case shows that the differences increase with shrinking buffer oxide thickness.

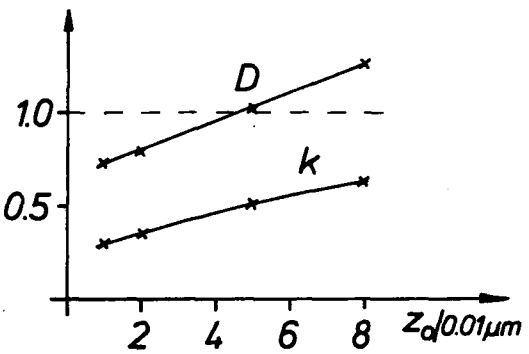


FIG. 5: Effective values for diffusion coefficient (D) and reaction rate (k) depending on buffer-oxide thickness normalized with respect to the stress-free values.

To quantitatively estimate the influence of stress on the diffusion- and reaction rates the bird's beaks were re-calculated with the values for these rates adapted such that the best fit for the geomtry of the under-diffusion-region was obtained. The values, which can be interpreted as mean-values for the whole under-diffusion-region, are shown by Fig. 5. The mask effect appears to be strong for a small pad oxide thickness. In this case the mask is strongly bent. If the values of the coefficients are translated into stress-values via Kao's model (1) it can be concluded that for thin pad oxides the mask exerts strong normal stress along the interface and that

compressive hydrostatic pressure is dominating throughout the under-diffusion-region. The mask effect becomes weaker for thick pad oxides. The mask is only slightly bent and seems to enhance the region of tensile pressure shown by numerical calculation. Nonetheless a reduced value of the reaction rate can still be observed.

7. Variation of Nitride Thickness

To investigate mechanical effects samples were fabricated with different nitride thicknesses at different temperatures. A direct deformation of the oxide, which leads to an assymetric profile could only observed for the maximum nitride thickness (2000Å) at 1100° (Fig. 6a). For the same nitride thickness, but lower temperature this effect is much less pronounced (Fig. 6b). This is due to the fact that the oxide becomes softer as compared to the nitride for high temperatures. In the second case the mask acts only indirectly through stress-induced reduction of the reaction rate. This results in stronger under-diffusion and thus leads to a very long bird's beak.



a



b

FIG. 6: Effect of nitride mask, 2000 Å: direct deformation at 1100 °C (a), influence via reduced reaction rate at 950° (b)

CONCLUSION

Extensive experimental data concerning the influence of technology parameters on the shape of MOS field oxide were obtained and characterized. Oxide growth is strongly influenced by mechanical effects. It was shown that the influence of stress on the reaction rate plays the main role. In addition the diffusion coefficient is enhanced significantly for some cases. Quantitative data on these effects were extracted.

ACKNOWLEDGEMENTS

The authors wish to thank B. Brandt, B. Schmiedt and E. Rose for performing the MOS-Technology. Moreover they are indepted to E. Lorenz, R. Schork, A. Daurer and K. Hartlieb for support of TEM microscopy.

REFERENCES

[1] R. B. Marcus, T. T. Sheng, J. Electrochem Soc., June 1982

[2] L. O. Wilson, J. Electrochem. Soc., Vol. 129, No. 4, pp. 831-837, April 1982

[3] T. C. Wu, W. T. Stacy, K. N. Ritz, J. Electrochem Soc., Vol. 130, No. 7, pp. 1563-1566, July 1983

[4] D. B. Kao, Stanford University, Technical Report No. G503-2, June 1986

VERIFICATION OF ION IMPLANTATION MODELS BY MONTE CARLO SIMULATIONS

G. Hobler, S. Selberherr

Institut für Allgemeine Elektrotechnik und Elektronik
Technical University of Vienna
Gußhausstraße 27-29, A-1040 Vienna, AUSTRIA

Monte Carlo simulations are perfectly suited to check the validity of simple models. We investigate 3 models: First, we show that 1D models for the implantation into multilayer targets give reasonable results only if the stopping powers of mask and bulk material are similar. Second, we discuss the construction of 2D point responses from 1D profiles. Third, we show that the method of superposing point responses at mask edges may fail in some cases.

1. INTRODUCTION

The Monte Carlo method is known to be the most powerful tool for the simulation of ion implantation. Analytical models, however, require much less CPU times and allow easy consideration of experimental data. The latter is particularly important because Monte Carlo simulations usually assume amorphous targets so that they do not always yield correct profiles for implantations into crystalline targets [1].

As simple models are usually based on physical considerations and Monte Carlo simulations take physics most accurately into account (apart from the assumption of amorphous targets), Monte Carlo simulations are perfectly suited to check the validity of these simple models. In particular, we will investigate in this paper 1D models for the implantation into multilayer targets (Chapter 2), the construction of 2D point responses from 1D profiles (Chapter 3), and the method of superposing point responses to obtain dopant distributions near mask edges (Chapter 4).

Our Monte Carlo program is, from a physical point of view, similar to the well known program TRIM [2]. One mayor difference of our code is that we evaluate scattering angles by interpolation in a precomputed table. The 2D simulations have been performed with a code which allows arbitrary geometries. Both features are described in Ref. [3].

2. IMPLANTATION INTO MULTILAYER TARGETS

In a recent paper [4], Ryssel discussed 5 models for the implantation into multilayer targets. These models consider 3 situations:

1) Implantation into bare material 1 (concentration profile $C_1(x)$).
2) Implantation into bare material 2 (concentration profile $C_2(x)$).
3) Implantation into a mask/bulk structure with given mask thickness d, where the mask material is material 1 and the bulk material is material 2 (concentration profile $C(x)$).

The purpose of the models is to construct $C(x)$ from $C_1(x)$ and/or $C_2(x)$. $C_1(x)$ and $C_2(x)$ may be obtained by simulations as well as by experiments. The models read:

$$C(x) = \begin{cases} C_1(x) & x < d \\ \alpha \cdot C_2\left(x - d \cdot \left(1 - \frac{R_{p2}}{R_{p1}}\right)\right) & x > d \end{cases} \quad (1)$$

$$C(x) = \begin{cases} C_1(x) & x < d \\ C_2\left(x - \left(d - d'\right)\right) & x > d \end{cases} \quad (2)$$

$$C(x) = \begin{cases} \frac{R_{p2}}{R_{p1}} \cdot C_2\left(\frac{R_{p2}}{R_{p1}} \cdot x\right) & x < d \\ C_2\left(x - d \cdot \left(1 - \frac{R_{p2}}{R_{p1}}\right)\right) & x > d \end{cases} \quad (3)$$

$$C(x) = \begin{cases} \frac{\Delta R_{p2}}{\Delta R_{p1}} \cdot C_2\left(\frac{\Delta R_{p2}}{\Delta R_{p1}} \cdot x\right) & x < d \\ C_2\left(x - d \cdot \left(1 - \frac{\Delta R_{p2}}{\Delta R_{p1}}\right)\right) & x > d \end{cases} \quad (4)$$

$$C(x) = \begin{cases} C_1(x) & x < d \\ \frac{\Delta R_{p1}}{\Delta R_{p2}} \cdot C_1\left(\frac{\Delta R_{p1}}{\Delta R_{p2}} \cdot x - \right. & x > d \\ \qquad \left. - d \cdot \left(\frac{\Delta R_{p1}}{\Delta R_{p2}} - 1\right)\right) & \end{cases} \quad (5)$$

α in (1) and d' in (2) are adjusted in such a way that $\int C(x)dx = \int C_1(x)dx$ $(= \int C_2(x)dx)$, what is automatically fulfilled in Models 3, 4, and 5. R_{p1}, R_{p2} denote the mean projected range and ΔR_{p1}, ΔR_{p2} the standard deviation of $C_1(x)$, $C_2(x)$.

Ryssel gave qualitative arguments in favour of Model 1. To investigate the models quantitatively, we have calculated $C_1(x)$, $C_2(x)$ and $C(x)$ by Monte Carlo simulations and then constructed $C(x)$ from $C_1(x)$ and $C_2(x)$ by applying one of the Models 1–5. Comparing the two versions of $C(x)$, one can easily see how good the models are.

Two examples are shown in Fig.1 and Fig.2. Fig.1 shows good agreement between Model 1 and Monte Carlo results for an As-implantation into SiO_2/Si. In Fig.2 can be seen, however, that the model fails completely for a Be- implantation into SiO_2/GaAs. In this case the profile in bare SiO_2 would describe the profile in SiO_2/GaAs much better than the profile constructed by Model 1. This indicates that the models fail, if mask and bulk material have very different stopping powers like SiO_2 and GaAs.

To confirm this result, we have performed simulations for B-, As-, Sb-, and Be-, Si-, Zn- implantations into SiO_2/Si and SiO_2/GaAs, respectively, at 3 different energies and for 3 values of the mask thickness. P-implantations have not been considered because P-profiles in SiO_2 and Si are almost identical. The energies are usually 30 keV, 100 keV, and 500 keV (10, 80, 500 for B and Be), the values for the mask thickness about $\frac{1}{5}R_p$ ("thin"), $\frac{4}{5}R_p$ ("medium"), $\frac{7}{5}R_p$ ("thick"). In order to present the results in a compact manner, we have introduced 4 degrees (cf. Tab.1 and Tab.2): "good" means that the profiles deviate in depth far less than 10%, "fair" means less than 10%, "poor" more than 10%. "catastrophic" has been introduced to indicate that one of $C_1(x)$, $C_2(x)$ would represent the profile in the mask/bulk structure better than $C(x)$ as calculated from the model.

In Tab.1 and Tab.2 there is listed for each mask thickness and each model the number of cases with good, fair, poor, and catastrophic agreement. (Note that the sum of each column is 9, as we have 3 ion species at 3 energies). In Tab.1, which is for SiO_2/Si, it can be seen that the general agreement is quite good, however, only Models 1 and 3 are always "good" or "fair", and Model 1 is slightly better than Model 3, in agreement with Ryssel [4]. On the other hand, all models completely fail for SiO_2/GaAs (Tab.2). Only for thin masks Model 3 gives good results.

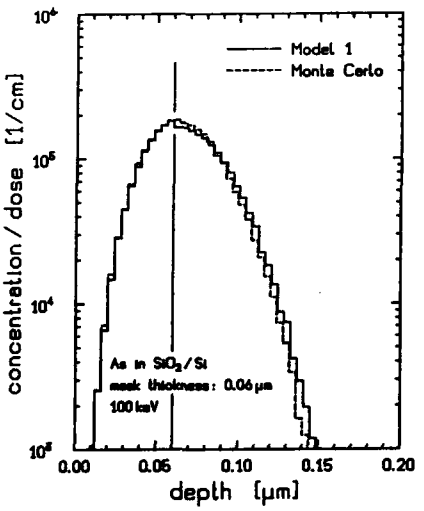


FIGURE 1

As-implantation into Si through a SiO_2 mask.
dashed line: Monte Carlo profile in SiO_2/Si.
full line: Profile in SiO_2/Si due to Model 1, constructed from Monte Carlo profiles in bare SiO_2 and bare Si.

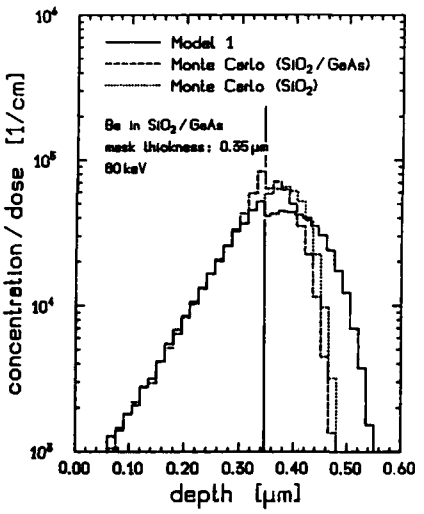


FIGURE 2

Be-implantation into GaAs through a SiO_2 mask.
dashed line: Monte Carlo profile in SiO_2/GaAs.
full line: Profile in SiO_2/GaAs due to Model 1, constructed from Monte Carlo profiles in bare SiO_2 and bare GaAs.
dotted line: Monte Carlo profile in bare SiO_2.

mask	thin					medium					thick				
model	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
good	9	4	9	7	1	6	5	5	3	5	9	9	4	1	9
fair	-	3	-	2	2	3	4	4	1	4	-	-	5	2	-
poor	-	1	-	-	6	-	-	-	5	-	-	-	-	6	-
catastrophic	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-

TABLE 1

Number of cases with good, fair, poor, and catastrophic agreement for implantations into SiO_2/Si.

mask	thin					medium					thick				
model	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
good	3	-	6	2	-	-	-	-	-	-	-	-	-	-	-
fair	2	-	2	-	-	3	-	1	-	-	-	-	-	-	-
poor	-	2	-	3	-	3	1	3	1	1	1	1	1	-	-
catastrophic	4	7	1	4	9	3	8	5	8	8	8	8	8	9	9

TABLE 2

Number of cases with good, fair, poor, and catastrophic agreement for implantations into SiO_2/GaAs.

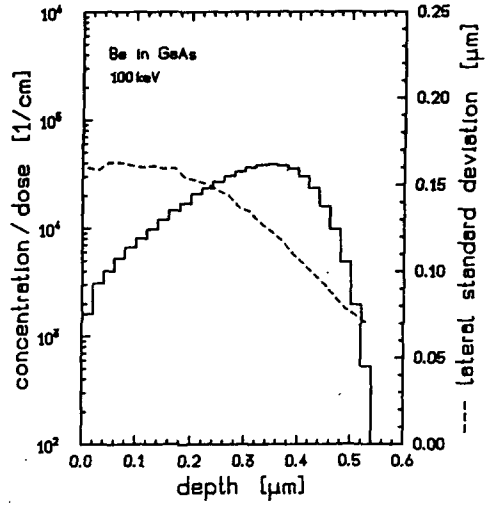


FIGURE 3

Depth dependence of the lateral standard deviation for Be in GaAs (100 keV).

3. CONSTRUCTION OF POINT RESPONSES FROM 1D PROFILES

Responses to punctiform beams play an important role in the Superposition Method (see Chapter 4). For a long time it was believed that one parameter, namely the lateral standard deviation, would be enough information to construct the 2D point response $C(x,y)$ from the 1D profile $C_{vert}(x)$. This was simply done by multiplying $C_{vert}(x)$ with the lateral Gaussian function $gauss(y)$ given by σ_y:

$$C(x,y) = C_{vert}(x) \cdot gauss(y) \qquad (6)$$

This means that the lateral profile at any depth is a Gaussian function with fixed standard deviation. In a previous paper [5] we have shown that this is not true for Si-targets. The lateral standard deviation depends strongly on the depth, and also the lateral profile is not always well represented by a Gaussian function.

We have now investigated GaAs-targets, and we found quite the same behaviour as for Si: For light ions (Be) the lateral standard deviation decreases with depth (Fig.3) and the lateral kurtosis is smaller than 3. For heavy ions (Zn) the standard deviation increases with depth and the kurtosis may assume large values near the surface. For Si-ions, which lie between the two cases, σ_y does not depend very much on the depth.

4. SUPERPOSITION METHOD

The superposition law says that the response to a homogenous beam is identical to the sum of responses to punctiform beams which are equidistributed over the width of the homogenous beam. For a rigorous application of this law we would have to know the actual response to every punctiform beam along the surface. In practice, however, point responses are constructed from 1D profiles and may therefore not take into account boundaries other than perpendicular to the beam. In the case of a mask edge those ions are not treated correctly by the superposition method which leave the mask laterally and re-enter the target. The question is now, whether these ions may significantly contribute to the total dopant concentration.

To investigate this question, we have performed Monte Carlo simulations for a simple structure, namely a rectangular mask on a planar bulk. In this case, according to the superposition method, no ions should reach the Si-region which have originally entered the mask. So, if we only expose the mask surface to the computational ion beam, any concentration in the Si-region indicates a failure of the model. We have performed simulations for B- and As-implantations at various energies. The results for B at 100 keV are shown in Fig.4. The concentration in Si is about one order of magnitude lower than the peek concentration of di-

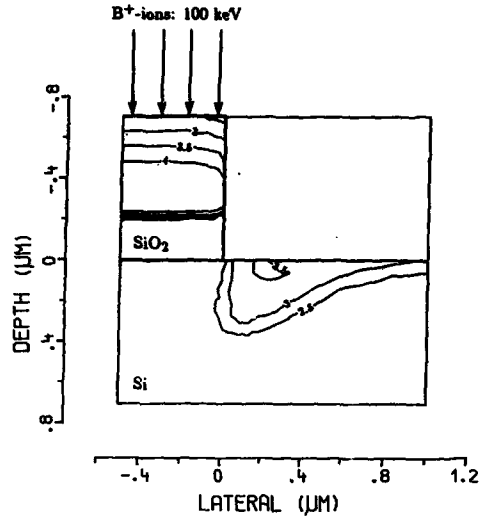


FIGURE 4

B-implantation into Si by a rectangular mask.
The contour lines represent the logarithm of the dopant concentration devided by the dose [1/cm].
Only the mask region is exposed to the beam.

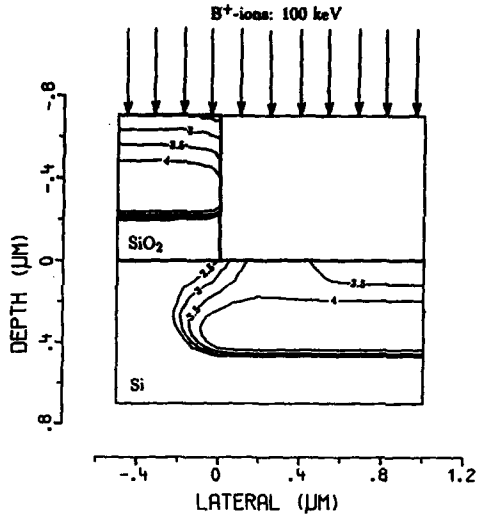


FIGURE 5

B-implantation into Si by a rectangular mask.
The contour lines represent the logarithm of the dopant concentration devided by the dose [1/cm].
The whole simulation area is exposed to the beam.

rectly implanted ions. This is typical for all cases we have simulated.

In Fig.5 it can be seen that this extra concentration —as compared with what is expected by the superposition method—contributes significantly to the total distribution. According to the superposition method, the contour lines labeled by "3.5" and "4" should be straight lines for lateral coordinates from slightly larger than 0 up to 1. Also the contour line labeled by "3" should be seen there.

For As- and low energy B-implantations this extra concentration may be well neglected, because in these cases the profiles have their maximum near the surface and will therefore cover the dopants which have made their way through the mask. A similar situation as in Fig.4 and Fig.5 is expected for high-energy P-implantations.

To avoid this effect, one could use a thicker mask, since the ions which leave the mask laterally will then spread over a wider range. E.g., for a mask thickness of $2\mu m$ in Fig.5 the effect would almost disappear. Another possibility would be to tilt the mask edge. In this case, however, the dopant distribution below the mask edge would be increased.

ACKNOWLEDGEMENTS

This work has been supported by the research laboratories of SIEMENS AG at Munich, by DIGITAL EQUIPMENT CORP. at Hudson, USA, and by the "Fonds zur Förderung der wissenschaftlichen Forschung", project S43/10.

REFERENCES

[1] H. Ryssel, J.P. Biersack:
"Ion Implantation Models for Process Simulation",
In: *Process and Device Modeling* (Ed. W.L. Engl), Elsevier, North Holland, pp. 31–69, 1986.

[2] J.P. Biersack, L.G. Haggmark:
"A Monte Carlo Computer Program for the Transport of Energetic Ions in Amorphous Targets"
Nucl. Instr. Meth., Vol. 174, pp. 257–269, 1980.

[3] G. Hobler, S. Selberherr:
"Efficient Two-Dimensional Monte Carlo Simulation of Ion Implantation"
Proc. NASECODE V Conf., Dublin, 1987.

[4] H. Ryssel, J. Lorenz, K. Hoffmann:
"Models for Implantation into Multilayer Targets"
Appl. Phys., Vol. A41, pp. 201–207, 1986.

[5] G. Hobler, E. Langer, S. Selberherr:
"Two-Dimensional Modeling of Ion Implantation with Spatial Moments"
Sol.-State Electron., Vol. 30, No. 4, pp. 445–455, 1987.

A SIMPLIFIED MODEL FOR THE CHARACTERIZATION OF ANTIMONY ION IMPLANTATION AND DIFFUSION ON SILICON

REUSI INÊS FONSECA

Laboratório de Microeletrônica da Escola Politécnica da USP
Departamento de Engenharia de Eletricidade, P.O. Box 8174
05508 São Paulo, Brasil

The main purpose of this work is to show how with a simple analytical model of diffusion, using only $R_\square$ and x_j experimental data, is it possible to calculate the relationship between the electrically active charge and the initial implanted dose of Antimony in Silicon. It will be shown that $Q_{el}/Q_{dose} < 1$ in agreement with many others authors[1,2,3], and a lower diffusion coefficient than that commonly used by SUPREM II simulator is achieved through this model, which allows good fit with the experimental data.

1. INTRODUCTION

High dose implantations of Antimony in Silicon studies are receiving considerable attention in recent years, by many authors[1,2,3,4], owing to their applications as an impurity source in the fabrication of buried-layers in high speed, low power dissipation Bipolar Transistors[5,6], and resistors with special characteristics[7], convenient for VLSI circuits. It is well known that for small concentrations of Sb in Si, almost 100% substitutionallity and electrical activity is achieved[2,3], whereas high concentrations exceeding the solid solubility limits of Sb in Si form metastable solutions and cause segregation effects, indicating that a fraction of substitutional Sb is electrically inactive. A complete characterization of these high concentrations of Sb in Si, their decomposition into precipitates, informations about the exact Sb atoms crystallographic location in the lattice and the degree of their electrical activity has been extensively studied, through a variety of analytical techniques[2,3].

This work shows through a simplified analytical model of diffusion, using only $R_\square$ and x_j experimental data, that is it possible to calculate the relationship between the electrically active charge, Q_{el}, and the initial implanted dose, Q_{dose}, of Sb in Si, together with the following parameters: diffusion coefficient D, mean mobility $\bar{\mu}$ and carrier concentration $\bar{C}$.

These results are in good agreement with those published earlier by many others authors, indicating that this model, although very simple, is sufficiently adequated to describe the fenomena correlated with the Sb diffusion on Si.

2. THE ANALYTICAL MODEL

After annealing and drive-in diffusion, impurity concentration redistribution from an ion-implanted source can be treated as a Gaussian profile, with fixed amount of impurities, given by Q_{dose}. Therefore, it is well known that only a fraction of these Q_{dose} is electrically active, named Q_{el}, and given by:

$$Q_{el} = \frac{1}{q\,\bar{\mu}\,R_\square} \qquad (1)$$

From $R_\square$ experimental data, the only way to calculate the actual Q_{el} is through the mean mobility $\bar{\mu}$ variations, which have a dependence over the mean impurity concentration $\bar{C}$. These parameters can be interconnected through the following definitions:

$$\bar{C} = \frac{Q_{el}}{x_j} \qquad (2)$$

and

$$\bar{\mu} = \mu min + \frac{\mu max - \mu min}{1 + \left(\frac{\bar{C}}{N_{ref}}\right)^{\alpha}} \qquad (3)$$

where:

$\mu min = 86{,}5\ cm^2\ V^{-1}\ s^{-1}$

$\mu max = 1354{,}5\ cm^2\ V^{-1}\ s^{-1}$

$\bar{C}$ = mean impurity concentration

$\alpha = 0{,}91$

$N_{ref} = 1{,}3 \times 10^{17}\ cm^{-3}$

It can be shown that for a long time diffusion the junction depth, x_j, is given by:

$$x_j^2 = 4Dt\ ln\ \frac{Q_{dose}}{C_B \sqrt{\pi D_i}} \qquad (4)$$

A linear function is obtained from $x_j \times \sqrt{t}$ plotting, whose gradient $dx_j/d\sqrt{t}$ allows the diffusion coefficient D determination from relation (4):

$$\sqrt{D} = \frac{dx_j}{d\sqrt{t}} \cdot \frac{1}{2\left[ln\ \frac{Q_{dose}}{C_B \sqrt{\pi D_i}}\right]^{1/2}} \qquad (5)$$

The first step in an iterative calculation between the relations (1), (2), (3) and (5), considers that all implanted impurity is electrically active, so that it is obtained the first values of D, $\bar{C}$, $\bar{\mu}$, and Q_{el}. Subsequent iteration calculations will allow to get the actual values of $\bar{D}$, $\bar{C}$, $\bar{\mu}$ and Q_{el}, when convergence is achieved.

As will be seen in the next item, using only the x_j and $R_\square$ experimental data it will be possible to obtain results comparable with those from another authors, whose experimental data were obtained from several technological facilities, like RBS analysis, Mössbauer spectroscopy, Hall-effect measurements, and so on[2,3].

3. EXPERIMENTAL PROCEDURES

All Sb implantation were performed in a home-made equipment, with an energy E = 100 Kev and dose $\phi = 5 \times 10^{15} cm^{-2}$, at room temperature, in a 7° off-axis direction, into silicon wafers type P, <100>, ρ = 10-20Ω.cm. After typical annealing at low temperature, T = 500°C, and O_2 ambient, for 60min, it was carried out the dopant diffusion in O_2 ambient, T = 1200°C, in 4 different times, 4, 9, 16 and 25 hours. Subsequent measurements of $R_\square$ and x_j of those samples were used to calculate the diffusion coeeficient, mean impurity concentration, mean mobility and electrically active charge by means of that iterative procedure suggested by the model described earlier in this paper.

4. RESULTS AND DISCUSSION

Figure 1 shows experimentally determined junction dephts, x_j, and sheet resistance mean values $R_\square$, obtained for several samples over a wide range of heat treatment times, compared with the simulated values produced by SUPREM II, only with the purpose of a more detailed overview about Sb diffusion on Si behavior:

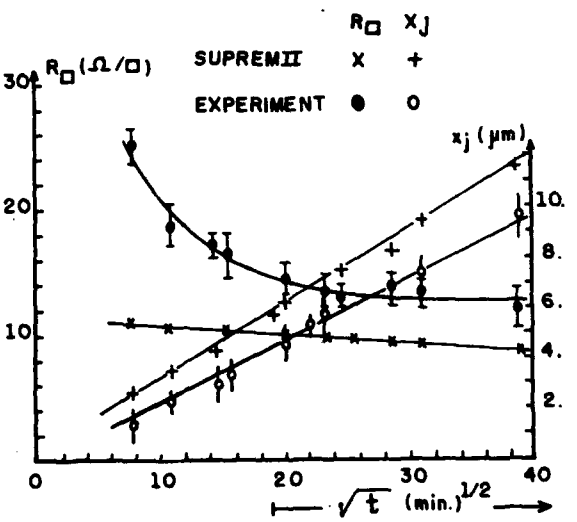


FIGURE 1

Comparison between SUPREM II and experimental $R_\square$ and x_j values.

As SUPREM II has not an Antimony cluster model, the predict data are overestimated in the sense that all implanted impurity is considered as electrically active, and the experimental results are below those simulated; moreover, the

saturation on $R_\square$ experimental data indicates a saturation on Q_{el}, due to the formation of extended defects like new Sb clusters or precipitates during the heat treatment[2], not considered by SUPREM II.

From $(dx_j/d\sqrt{t})_{EXP}$ of figure 1 and using the analytical model proposed on this work, it was calculated the new diffusion coefficient D value ($D \cong 1{,}16 \times 10^{-11} cm^2.min^{-1}$), and consequently, the $\bar{C}$, $\bar{\mu}$ and Q_{el} parameters. Modyfying SUPREM II by introducing the new values of D and Q_{el} as the initial implanted dose, it was obtained a good fit with experimental $R_\square$ and x_j values, as illustrates figure 2 below:

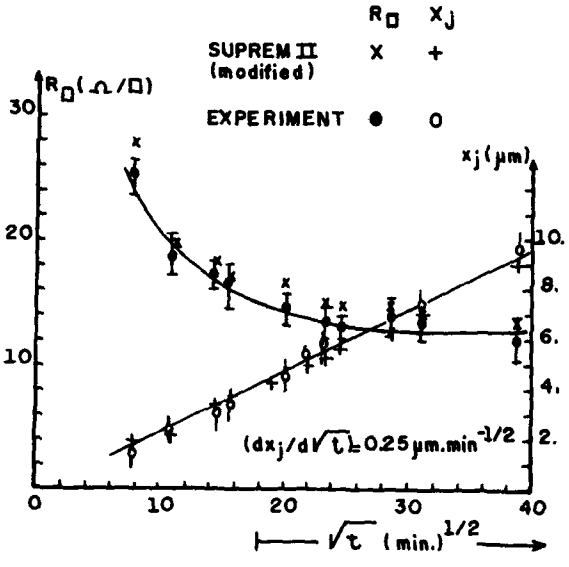


FIGURE 2

Comparison between modified SUPREM II and $R_\square$ and x_j experimental values.

Finally, on figure 3 it is shown the relation $Q_{el}/Q_{dose} \times \sqrt{t}$, where two points must be emphasized:

- $Q_{el}/Q_{dose} < 1$ indicates that only a fraction of implanted dose is electrically active, probably that one on undisturbed substitutional lattice sites. This percentual result shown in figure 3 is in good agreement with others authors[1,2,3];
- the saturation of Q_{el}/Q_{dose} for long time diffusion is an expected result, and indicates a reduction of the electrically active fraction of Antimony, probably by the formation of precipitates[2] or Sb-vacancy complexes[2,4].

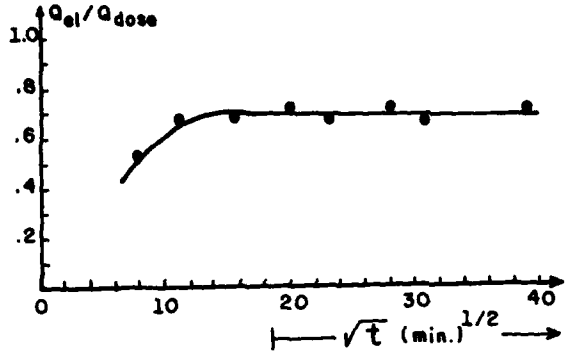


FIGURE 3

Q_{el}/Q_{dose} as calculed by the model proposed on this work.

Although not presented here the mean concentration, $\bar{C}$, and mean mobility, $\bar{\mu}$, as calculated by the model[6], seems to be in agreement with results from another authors[2,3,8].

5. CONCLUSIONS

In this work it was presented a very simple analytical model, that allows to get informations about the electrically active fraction of Antimony implanted on Silicon, taking into account only $R_\square$ and x_j experimental data. The results of the model have shown that the diffusion coefficient of Antimony on Silicon is lower than that used by SUPREM II simulator, and that only a fraction of the initial implanted dose is electrically active, as previously published by many others authors. The modification of SUPREM II by introducing these new data allowed a good fit with experimental $R_\square$ and x_j data.

ACKNOWLEDGEMENTS

The author would like to thank the students H. Peres and E. Galeazzo, for making the ion implantation and the experimental measurements, respectively, and Mrs. M. Brito for typing this manuscript. Helpful suggestions of Dr. S.Solmi are also acknowledged.

The partial financial support by FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo,is gratefully acknowledged by the author.

REFERENCES

1. JOSQUIN, W.J.M.J. and TAMMINGA, Y., Applied Physics, 15, (1978), 73-78.

2. NYLANDSTED LARSEN, A., et al, MRS Europe, (1985), 319-324.

3. NYLANDSTED LARSEN, A., et al, J. Appl. Phys., 59 (6), (1986), 1910-1917.

4. FAIR, R.B., et al, J. Mater. Res., 1 (5), (1986), 705-711.

5. TANG, D.D., et al, IEEE Trans. Electron.Dev., 27 (8), (1980), 1379-1384.

6. FONSECA, R.I., Ph.D. Thesis, Escola Politécnica, USP, (1985).

7. KU, S.M., and CHU, W.K., Solid-State Electronics, 22, (1979), 719-722.

8. JOHANSSON, N.G.E., and MAYER, J.W., Solid-State Electronics, 13 (1970), 123-130.

GLASS REFLOW MODELING FOR PROCESS OPTIMIZATION

A. TISSIER, A. PONCET and J.F. TEISSIER
CNET-Grenoble - France

1 INTRODUCTION

PSG and BPSG are intensively used in VLSI processes for their flow capability . In a micronic multilevel metallisation technology, it is necessary to control the flow annealing which tends to smooth the topology, particularly in two places : the gate overlap and the contact window steps . In the literature, work has been mentionned which deals essentially with measurements on SEM views of the tangential angle of the layer at the step edge as a funtion of the annealing parameters and the glass composition [1],[5],[7]. A new approach is presented here, which combines experimental results with numerical simulations of glass reflow, in order to predict the "optimal" annealing, i.e. an increased planarity and a minimization of parasitic thermally activated phenomena (dopant diffusion). Coupling SEM measurements and numerical simulations allows to process only one test pattern and, furthermore, to extrapolate the results to any case.

2 PHENOMENOLOGICAL STUDY OF THE VISCOUS FLOW

In order to minimize the induced technological dispersions, the simplest test pattern is chosen, i.e. a rectangular glass slab (0.8 μm height and 4 μm width). After cleaving , the geometrical evolution of cross sections is studied by SEM measurements as a fonction of the RTA parameters (T from 950 o C to 1190 o C , t from 10 to 80 s) and the glass composition (6% w/o P to 8.9% w/o P PSG and 5% w/o B, 5% w/o P BPSG). As depicted on figure 1 the glass reflow leads to the modification of the following three geometrical parameters

- the angle, θ,
- the thickness at the middle of the step, h,
- the curvature radius, R;

The first parameter being the more sensible to the reflow annealing, it is chosen to quantify the viscous deformation.

Experiments made on both gate overlap and test pattern show that for θ equal to 15 degrees, the planarisation is acceptable. We observe that even in simplest cases, it is difficult to directly compare experimental and simulated profiles because of the dispersion, wafer to wafer or run to run, on data related to the slab formation, i.e. thickness, CD, angle after etching and local glass composition, therefore, it is necessary to average data.

The evolution of θ according to the time is measured for different temperatures , and for various glasses. Let t denote the annealing time which leads to θ=15 degrees. t values are extracted from these measurements; next, ln(t) is plotted as a function of 1/T (figure 2).

From these curves it can be observed that :

- the linearity of this function allows to fit physical parameters related to glass viscosity by carriing out linear regressions (see section 3);
- moreover, when these curves are superimposed with similar curves related to other thermally activated phenomenon (here boron diffusion) it is possible to identify the optimal temperature range for a given glass.

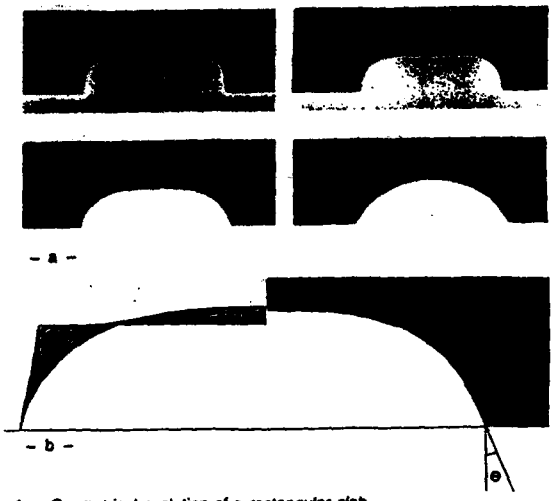


Fig. 1 – Geometrical evolution of a rectangular slab
– a – SEM views (8.6% w/o P PSG, T = 1190 C, t = 5, 10, 20 and 40 s)
– b – Simulation results compared with SEM view at t = 20 s.

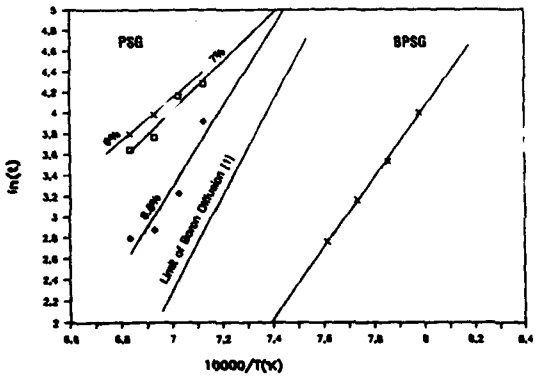


Fig. 2 – Logarithm of the time necessary to get Θ = 15° as a function of 1/T

3 NUMERICAL SIMULATION OF GLASS REFLOW

3.1 Viscous Flow Model

Under usual processing temperatures, the viscosity of passivation glasses is sufficiently high to assume that stationnary Stokes equations are valid for modeling their reflow. The driving force is the surface tension [8], which tends to smooth the free surface by increasing curvature radius in such a way that the cross section of any bounded slab of glass tends to be a piece of perfect disk (figure 1). However, such final shapes have no practical interest and are not valid for unbounded slabs, moreover, intermediate stages can not be accurately predicted under geometrical considerations only [4]. Therefore, computer simulations are necessary, i.e. discretization of Stokes equations. These equations can be summarized as follows in the 2-D case:

$$(3.1) \quad v\ \mathrm{Div}\ (\ \mathrm{Grad}\ \vec{V}\ ) = -\ \mathrm{Grad}\ p$$

and,

$$(3.2) \quad \mathrm{Div}\ \vec{V} = 0 \qquad \text{in the material,}$$

where $\vec{V}=(Vx,Vy)$ is the local velocity, v is the viscosity and p the internal pressure;

(3.3) $Vx = Vy = 0$ (non-slip condition) on the interface between glass and substrate,

(3.4) $Vx = 0$ (slip condition) along symetry axes and lateral sections,

(3.5) $p = Y / R$ along the free surface, where Y is a surface tension coefficient and R is the curvature radius.

A major application of glass reflow simulation concerns contact holes; for that purpose, an axisymetric expression of the equations has been set under a variational form; 3-D effects have been clearly evidenced in numerical experiments: Θ becomes much smaller when the radius of a contact hole decreases (figure 3).

3.2 Viscosity Fitting

The main advantage of the above model is its linearity according to v/Y; this ratio can be easily identified from experiments as follows:

1. arbitrary v/Y ratio and time scale are chosen (let say v/Y=1), then computer simulation is performed;

2. numerical results are compared with measurements (figure 1) in order to set the time scale which corresponds to a given glass: let tn be the time necessary to reach a given D value in the computer simulation, and tm the corresponding time deduced from experiments; therefore, the actual value of v/Y is tm/tn.

According to the two linearities mentionned above, relation (2.1) can be re-written, first:

$$(3.6)\quad \ln(t) = A + B/T$$

and then

$$(3.7)\quad \ln(v/Y) = A + B/T - \ln(tn)$$

which confirms the classical expression [2]:

$$(3.8)\quad v/Y = u0 \cdot \exp(E/kT)$$

while giving an straightforward evaluation of u0 and E parameters:

$$(3.9)\quad u0 = \exp(A)/tn$$

$$(3.10)\quad E = k \cdot B$$ where k is the Boltzmann constant.

By using measurements depicted on figure 1 for a 4 µm long and 0.8 µm high step, this method leads to values for E and u0 which are presented on table 1; however, the reproducibility of the slab dimensions and of RTA parameters, the accuracy of measurements and the temperature range are too low to quantify the dependence of E and u0 versus glass composition.

3.3 Numerical Schemes And Computer Environment.

Equations (3.1)-(3.5) are discretized by using classical 3-node triangular finite elements, meshes are automatically generated and refreshed, in the same way as in LOCOS simulation [7]. Incompressibility condition (3.2) is taken into account iteratively, by using classical Uzawa algorithm. Surface condition (3.5) is expressed through a boundary integral in the variationnal form of the equations.

Impurity diffusion and glass reflow have been coupled in TITAN process simulator [3], in order to achieve technological parameter optimization which has been mentionned in Section 1.

4 CONCLUSION

A simple linear viscous flow model has been presented in order to predict PSG or BPSG glass reflow. A method has been presented for identifiing viscosity parameters for any given glass, in order to optimize glass reflow, anywhere on the wafer.

However, the application of this approach depends drastically on the initial structure (composition, shape of the slab,...) and on RTA parameters variations; therefore, a general expression of viscosity versus temperature and glass composition cannot be set as long as these data are not accurate enough.

	PSG			BPSG
	6%	7%	8.6%	
E (eV)	5.15	4.56	1.25	1.08
u0	10^{-13}	10^{-11}	10^{-29}	10^{-28}

TABLE 1. Viscosity parameters from fits between measurements (Fig. 1) and computer simulation

5 REFERENCES

[1] N.S. ALVI and D.L. KWONG "Reflow of PSG by Rapid Thermal Annealing" Symp. on reduced temperature processing for VLSI, Proceedings Vol 86-5, edited by the Electrochemical Soc. (1986).

[2] D. CHIN "Two-Dimensional Oxidation, Modeling and Applications" PhD. Stanford, June 1983

[3] A. GERODOLLE, S. MARTIN and A. MARROCCO, "Finite Element Method Applied to 2-D MOS Process Simulation and defect diffusion: Program TITAN", NASECODE IV Conf. Proceedings, Boole Press, June 1985.

[4] R.A. LEVY and K. NASSAU " Reflow Mechanism of Contact Vias in VLSI Processing" J. Electrochem. Soc., Vol. 133 No. 7, pp. 1417-1424 (1986).

[5] J.S. MERCIER, R.P. BEERKERNS, I.D. CALDER and H.M. NAGUIB, Electrochemical Soc. ext. abstracts No. 420 Vol. 84-2 p. 607 (1984).

[6] A. PONCET, "Finite Element Simulation of Local Oxidation of Silicon", IEEE Trans. on Computer-Aided Design, Vol. CAD-4 No.1, pp. 41-53, 1985.

[7] T.O. SEDGWICK, F.M. D'HEURLE and S.A. COHEN, J. Electrochem. Soc. Vol. 131, p. 2446 (1984).

[8] P. SUTARDJA, Y. SHACHAM-DIAMAND and W.G. OLDHAM, "Two-Dimensional Simulation of Glass Reflow and Silicon Oxidation" I.E.D.M. Conf. , Los Angeles, Dec. 1986.

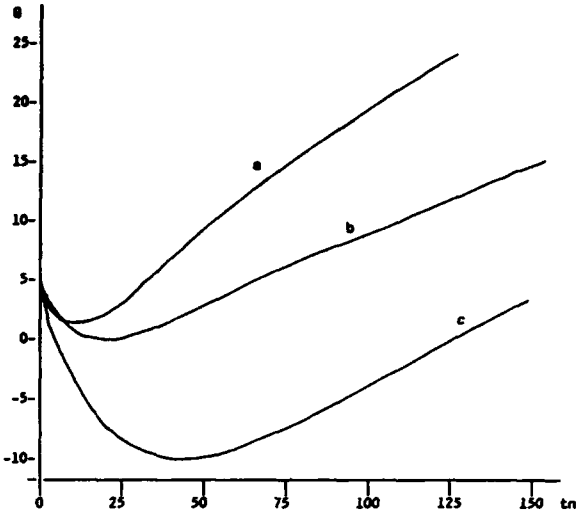


Figure 3. θ-angle versus time:
-a- 0.4 μm thick and 4 um long slab,
-b- 0.8 μm thick and 4 um long slab,
-c- cylindrical contact hole,
hight=0.4 um, radius=0.25 um.

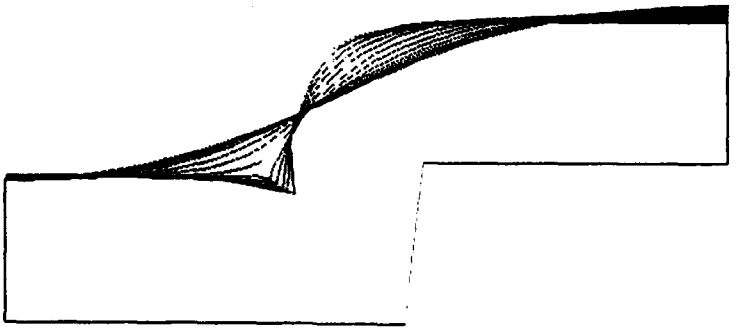


Figure 4. Numerical simulation of glass reflow at gate overlap

MONTE-CARLO ION IMPLANTATION AND COMPOSITE

A.Barthel, J.Lorenz, H.Ryssel*

Fraunhofer-Arbeitsgruppe für Integrierte Schaltungen,
Artilleriestrasse 12, D-8520 Erlangen, Germany

Analytical methods for the description of ion implantation show good agreement with experiment and Monte-Carlo simulations in most cases. Problems arise with special geometries such as trenches. To be able to simulate implantation and diffusion in such cases, a Monte-Carlo interface has been added to the process simulation program COMPOSITE.

1. INTRODUCTION

To meet the needs of shrinking device dimensions, process simulation programs are required which use accurate physical models for the simulation of process steps, use efficient algorithms to reduce computing time and are able as well to deal with a complete process sequence as to transfer the results as input for device simulation. These three requirements are very hard to be fulfilled with one simulation tool, as accurate process models very often require large computing time, for instance in case of Boltzmann transport equation calculations or Monte-Carlo simulations [1].

In the following, the approach to ion implantation used in COMPOSITE [2] is briefly mentioned along with its limitations. The Monte-Carlo interface which has been added to COMPOSITE is described and its application is shown.

2. COMPOSITE

The universal two-dimensional process simulation program COMPOSITE (Complete Modeling Program of Silicon Technology) is a user-friendly and easily-portable tool for the simulation of ion implantation, diffusion, oxidation, etching, lithography and layer deposition. For the simulation of ion implantation, analytical equations are used for the dopant concentration profiles. This includes the well-known Pearson IV-distributions [3] along with range parameters from experiments for the vertical dopant concentration profile in one layer, a lateral convolution with a Gaussian profile and a special multilayer model [4], which takes into account the different stopping powers of the layers. In figure 1, results obtained with COMPOSITE for an implantation of 60 keV phosphorus at an Al_2O_3-mask edge are compared to results obtained using the widely used Runge model [5] (broken lines), which assumes the same stopping power for all layers. According to the Runge model, the Al_2O_3 layer would not be thick enough to mask the silicon. From the COMPOSITE-result it can be seen that the Al_2O_3 thickness is sufficient to stop the ions.

3. MONTE-CARLO SIMULATIONS

Other methods for the simulation of ion implantation such as Monte-Carlo simulations [1] require much more computing time in comparison to analytical models and are, therefore, not suited for permanent use in a process simulation tool. But they are very important for the evaluation of analytical

*also: Lehrstuhl für Elektronische Bauelemente,
Universität Erlangen-Nürnberg,
Artilleriestrasse 12, D-8520 Erlangen, Germany

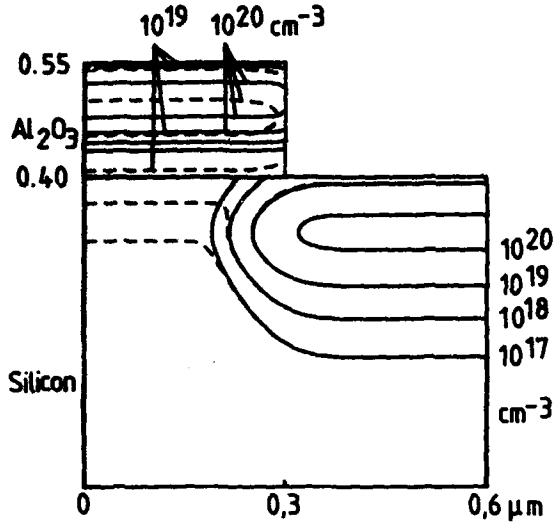


FIGURE 1

COMPOSITE-simulation of implantation of a 10^{15} cm^{-3} dose of phosphorus at an energy of 60 keV near an Al_2O_3 mask edge.

descriptions of implantation profiles and for the simulation of implantations into geometries which cannot be described adequately by analytical models. One main point of interest is the implantation into trenches in silicon.

Figure 2 shows the result from a Monte-Carlo simulation with a modification of TRIM, TRIMSURF [6], of an implantation of 150 keV phosphorus into a 2 µm deep and 0.4 µm wide silicon trench. 200 000 particles were used for this simulation. In this example, a sidewall-doping by ions which have been scattered out of one sidewall and have been re-implanted into the other sidewall can be seen. This sidewall doping is of great importance and cannot be accessed by analytical models. Therefore, it is very important to use such Monte-Carlo results within general simulation tools.

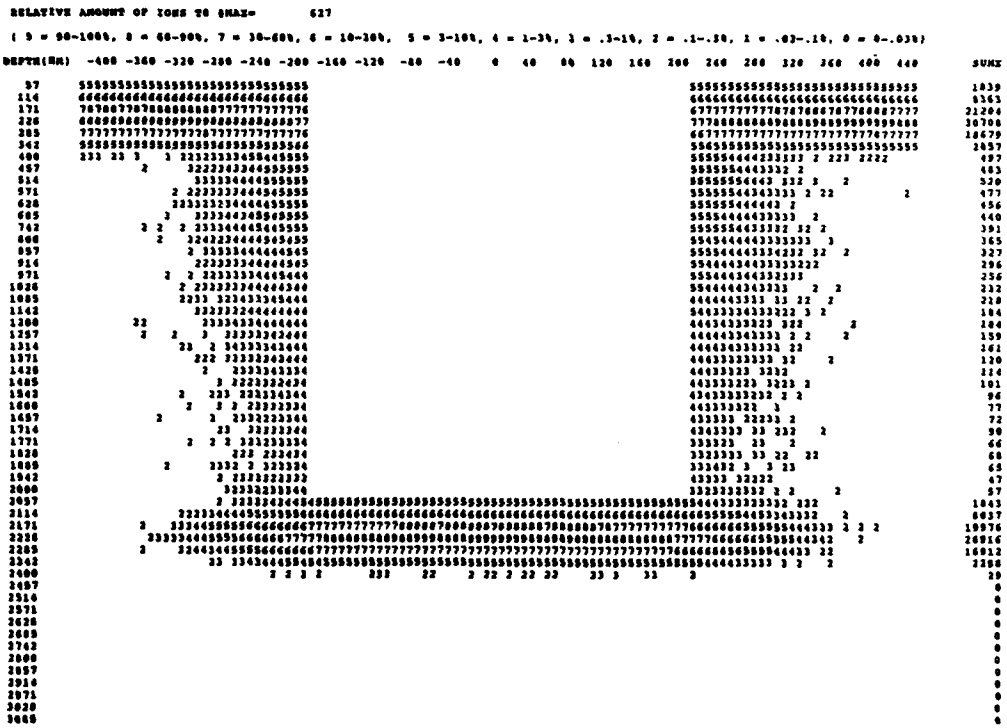


FIGURE 2

Monte-Carlo simulation of implantation of phosphorus at an energy of 150 keV into a 2µm deep and 0.4 µm wide silicon trench. 200 000 particles were used for this simulation.

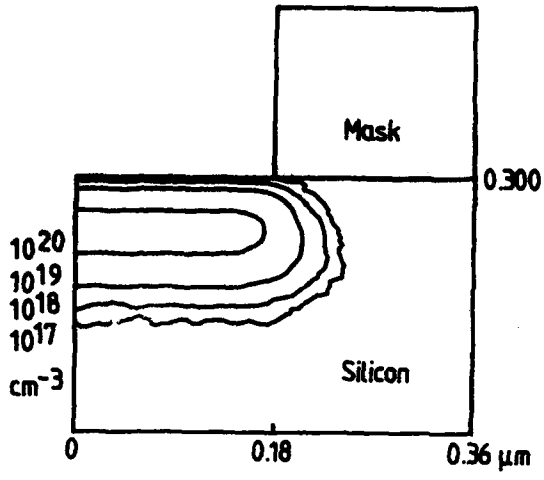


FIGURE 3

Monte-Carlo simulation of implantation of arsenic at an energy of 100 keV into silicon near a mask edge.

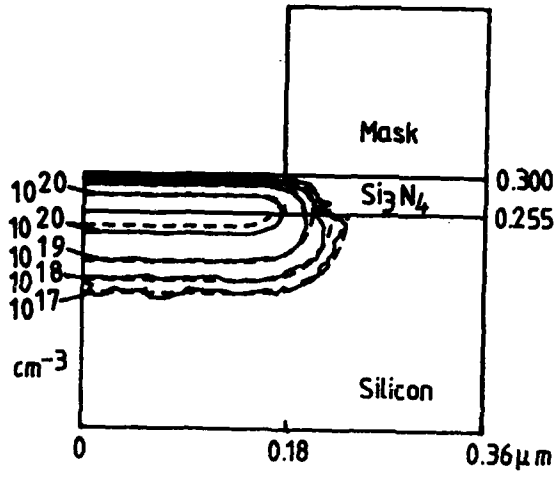


FIGURE 4

Comparison of TRIMSURF and COMPOSITE results for implantation of a dose of 10^{15} cm^{-3} arsenic at an energy of 100 keV into a two-layer structure near a mask edge.
Drawn line: TRIMSURF, broken line: COMPOSITE

4. USE OF MONTE CARLO SIMULATIONS FOR COMPOSITE

To be able to transfer results from Monte-Carlo simulations to COMPOSITE, an interface has been implemented.

First, some modifications to TRIMSURF have been done. These includes the gathering of the particle distribution data in a COMPOSITE-compatible shape. Furthermore, smoothing by neighborhood averaging is done to reduce statistic fluctuations with the data: Concentrations are recalculated as arithmetic means of the point in question and its eight nearest neighbors. The dopant concentration arrays are then stored to a file.

Second, COMPOSITE reads these data from the file and scales them according to the implantation dose desired.

In figure 3, an example for a Monte-Carlo simulation of an ion implantation of 100 keV arsenic into a silicon layer near a mask edge is shown. The Monte-Carlo data have been transferred to COMPOSITE. In the equiconcentration line plot the fluctuations in the third and fourth contour line result from the limited number of particles used with the Monte Carlo calculations. The maximum of the lateral spread of the ions implanted does not coincide with the maximum of the vertical distribution. This indicates the depth dependence of the lateral straggling, studied in earlier publications [7,8,9]. Equations for this depth dependence have been proposed [8], but they presently cannot be used for the simulation of ion implantation in crystalline silicon, because they need not only vertical moments but also lateral and mixed range moments, in total 8 parameters. The lateral kurtosis and the two mixed moments requested have not yet been measured or calculated for crystalline silicon. Therefore, the depth dependence is not included in COMPOSITE. For amorphous silicon, this model shows good agreement with Monte-Carlo simulation [8].

In contrast to figure 3, the silicon is covered by 45 nm Si_3N_4 in the example shown in figure 4. This is done to show the influence of a thin layer on the lateral spread in the silicon substrate. Since the lateral straggling in the nitride is smaller than in silicon because of the higher density of nitride, the

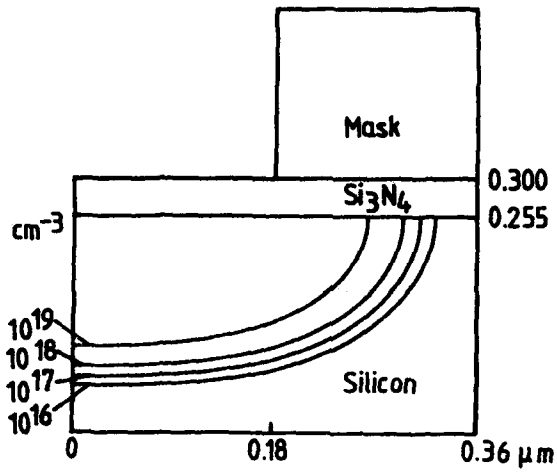


FIGURE 5

30 min diffusion at 1000 °C of the Monte-Carlo result shown in figure 4.

lateral spread of the implantation profile in the silicon is reduced in comparison to figure 3. Therefore, a multilayer model for the depth dependency of the lateral spread is necessary.

The broken lines in figure 4 show the corresponding COMPOSITE-results. Here, a constant lateral straggling was used within one material. Therefore, a discontinuity of the lateral spread is predicted by COMPOSITE, but TRIMSURF shows a nearly continuous behavior of the lateral spread at the interface. For the vertical dopant distribution, the agreement between COMPOSITE and Monte-Carlo is very good, except close to the silicon/nitride interface. The discontinuity of the vertical distribution at the interface results from the lower stopping power of the silicon: Therefore, less particles come to rest below the interface. This effect is less pronounced with the Monte Carlo data because of particles backscattered.

This example shows that as long as no full set of eight range parameters for the materials involved, including crystalline silicon, is available, no accurate simulation of the lateral spread is possible with analytical models. If the differences between amorphous and crystalline range parameters can be neglected, it is worthwhile to transfer Monte-Carlo data to COMPOSITE to perform the simulation of further process steps.

In figure 5, the COMPOSITE result of a 30 min diffusion at 1000°C of the doping profile from figure 4 is shown. The statistic fluctuations present in figure 4 have been removed by the diffusion.

5. CONCLUSION

Though the analytical equations for ion implantation used in COMPOSITE are able to describe dopant profiles adequately in most cases, the simulation of important effects such as trench implantation and depth dependence of the lateral spread presently needs time-consuming Monte-Carlo calculations. The Monte-Carlo interface implemented in COMPOSITE allows now for introducing Monte-Carlo data into a process sequence.

REFERENCES

[1] Biersack, J.P. and Haggmark, L.G., Nucl.Instrum. Methods 174 (1980) 257

[2] Lorenz, J., Pelka, J., Ryssel, H., Sachs, A., Seidl, A. and Svoboda, M., IEEE Trans. El. Dev. ED-32 (1985) 1977

[3] Hofker, W.K., Philips Res. Repts., Suppl. No. 8 (1975)

[4] Ryssel, H., Lorenz, J. and Hoffmann, K., Appl. Phys. A 41 (1986) 201

[5] Runge, H., Phys. Stat. Sol. (A) 39 (1977) 595

[6] Ryssel, H., Lorenz, J. and Krüger, W., Nucl. Instrum. Methods B 19/20 (1987) 45

[7] Hobler, G., Langer, E. and Selberherr, S., Two-dimensional modelling of Ion-Implantation, in: Board, K., Owen, D.R.J. (ed.), Simulation of Semiconductor Devices and Processes, Vol 2, (Pineridge Press, Swansea, U.K., 1986) pp. 256-270

[8] H. Ryssel et al., 4th report on "Zweidimensionale Prozeßsimulation vollständiger technologischer Prozeßabläufe", Erlangen, 1986

[9] Ashworth, D.G. and Oven, R., Computer simulation of the lateral spreading of implanted ions, ESSDERC '86

EQUILIBRIUM SOLUBILITY OF ARSENIC AND ANTIMONY IN SILICON

R.ANGELUCCI, A.ARMIGLIATO, E.LANDI, D.NOBILI, S.SOLMI

CNR - Istituto LAMEL, Via Castagnoli 1, 40126 Bologna, Italy

Equilibrium solid solubility of arsenic and antimony in silicon is derived by Hall and resistivity measurements after suitable annealing. For both elements, the solubility shows a linear trend versus reciprocal temperature.

1. INTRODUCTION

The knowledge of solid solubility of dopants in silicon is essential for a correct process simulation and is important for basic understanding In the case of Group V dopants largely scattered values are reported in literature for arsenic and antimony. Moreover, the knowledge is even poorer in the range 700÷900°C, which is of high interest in the future VLSI-ULSI processing.

A research activity on the solubility and precipitation of silicon dopants is performed since several years at LAMEL Institute. Particular emphasis was given to the study of electrically inactive phosphorus and arsenic; an assessment of this problem was attempted by Nobili a few years ago /1/.

This paper reports the results of accurate equilibrium carrier density determinations as a function of temperature, performed on polysilicon films heavily doped with antimony and arsenic by ion implantation. The carrier density was determined after annealing at increasing temperatures, a time consuming procedure which, on the other side, is most suitable to accomplish with the equilibrium conditions.

2. EXPERIMENTAL

Poly-silicon films were deposited in a chemical vapour reactor at 660°C, onto previously oxidized single crystal wafers. The film thickness (0.45 μm) was accurately determined by a Taylor Hobson Talystep.

For Sb doped films the implantation energy and dose were 150 keV and 2.1×10^{16} at/cm^2 respectively, while As was implanted at the energy of 100 keV, and dose 4.0×10^{16} at/cm^2. Each composition was then separately heated 3 hours at 1100°C to recover the damage and redistribute the dopant. Specimens were successively annealed for 1000 h at 600°C, then at temperatures incresing in steps of 25°C up to 900°C and subsequently in steps of 50°C up to 1300°C. To avoid out diffusion processes the first high temperature heat treatment was performed in a slightly oxidizing atmosphere (90% nitrogen + 10% oxygen).

The carrier density and mobility were determined by Hall effect and sheet resistivity measurements, using the Van der Pauw geometry defined with a photolitographic process. The same techniques, alternated with stripping of silicon by anodic oxidation and etching, were used for carrier profile measurements. The average grain size of the polysilicon films, after the high temperature annealing (1100°C, 3h) was checked by transmission electron microscopy observations. The obtained values were 0.9 μm and 0.2 μm for As and Sb, respectively.

3. RESULTS AND DISCUSSION

The carrier density plot vs reciprocal tempe-

rature for Sb doped specimens, which is reported in Fig.1, shows an initial decrease which can be attributed, as it is discussed below, to the formation of the conjugate liquid phase. A minimum is attained at 800°C followed by dissolution which takes place with increasing temperature. Equilibrium values of the carrier density n_e are obtained in the dissolution stage, above 850°C, after a transient which is due to the size effect. The equilibrium values of n_e in the temperature range 850-1150°C follow very tightly the law:

$$n_e = 3.8 \times 10^{21} \exp(-0.56 \text{ eV}/kT) \quad \text{cm}^{-3}$$

Above 1150°C the experimental n_e values show a deviation from this trend. This was expected as in the Si-Sb equilibrium diagram the conjugate liquid phase undergoes a drastic reduction of the content of antimony /2/.

The results obtained on As doped specimens are shown in Fig.2, which reports the carrier concentration as a function of reciprocal temperature. In this case, due to the higher diffusivity of arsenic with respect to antimony, the minimum was attained at a lower temperature, about 650°C, and equilibrium values of the active dopant were obtained for $T \geq 700$°C. In the temperature range 700-900°C the corresponding equilibrium carrier density n_e is given by:

$$n_e = 2.2 \times 10^{22} \exp(-0.47 \text{ eV}/kT) \quad \text{cm}^{-3}$$

These figures coincide with the ones reported by Hoyt et al. /3/, obtained by a fitting of literature data. A deviation from the above trend is observed in Fig.2 at higher temperatures, a result which is unexpected considering that the eutectic temperature is 1097°C /4/. We point out that the value at 1100°C in Fig.2 coincides with the one after the initial 3 hours annealing at this temperature. The above equilibrium values of

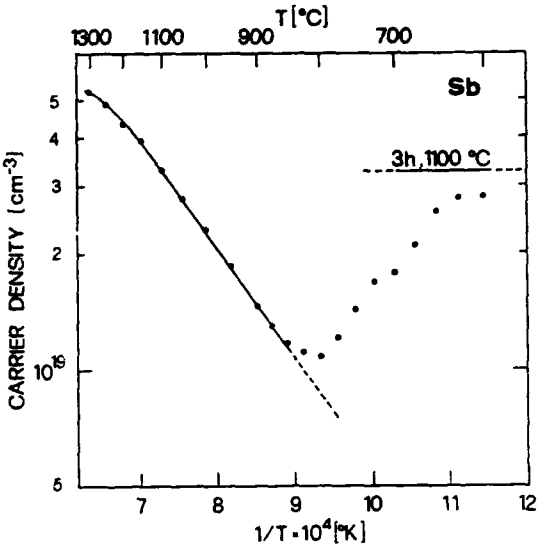


FIGURE 1

Carrier concentration vs reciprocal temperatures for Sb doped samples.

Arsenic are in very good agreement with the ones obtained by carrier profiles measurements after equilibration annealing of single crystal specimens implanted with different doses of the dopant /5/. These experiments

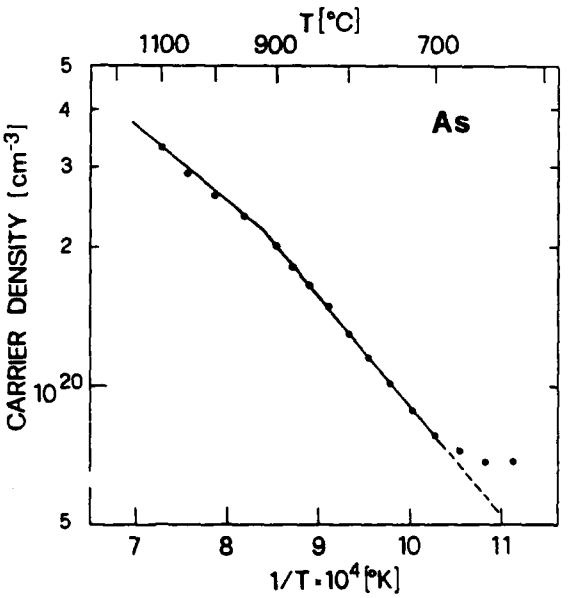


FIGURE 2

Carrier concentration vs reciprocal temperature for As doped samples.

showed that the carrier density after thermal equilibration depends only on temperature and is insensitive to excess dopant. We concluded from this results, which are supported by the occurrence of reversion and by TEM and SAXS examinations, that the equilibrium carrier density corresponds to the solubility. The same conclusion was reached also in the case of antimony by additional experiments performed on the same line, i.e. accurate carrier profile measurements after equilibration at 1100°C of single crystal specimens implanted at 160 keV with three different doses of the dopant. The results are shown in Fig.3 for an annealing time of 4 h.

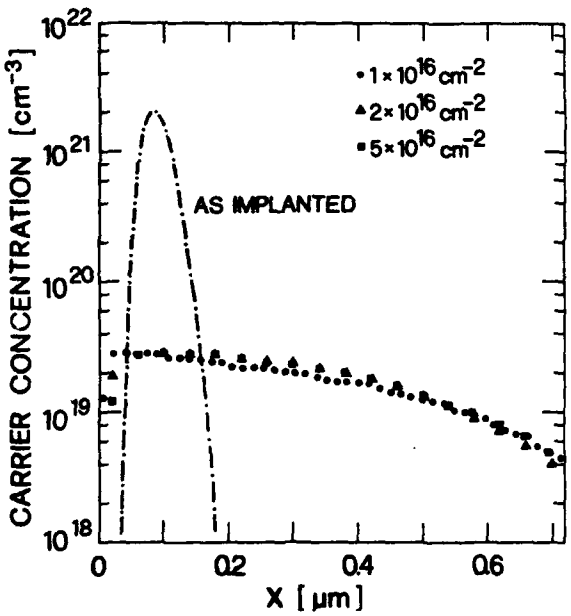


FIGURE 3

Carrier concentration profiles of Sb doped specimens implanted at different doses and annealed at 1100°C for 4 h. The as-implanted distribution for the lowest dose is also reported.

TEM observations performed on these samples evidenced the presence of a high density of Sb particles, having a size which decreases with increasing the implanted dose. According to the classical nucleation theory /7/, the density of the precipitates increased by increasing the supersaturation. In Fig.4 is reported a dark-field image of these particles, taken in a sample implanted with $2x10^{16}$ Sb/cm^2. They have the structure of the hexagonal antimony, as deduced from electron diffraction patterns.



FIGURE 4

Dark-field TEM micrograph, showing Sb precipitates in a $2x10^{16}$ cm^{-2} implanted sample, annealed at 1100°C for 4h.

High temperature data, above 1100°C, for As are not reported in Fig.2 because we verified that they were affected by the cooling rate. This phenomenon, which is attributed to additional precipitation taking place in the cooling stage, is more effective in polycristalline specimens. In fact dislocations and grain boundaries enhance the diffusion and nucleation kinetics of the dopant, a feature which, on the other side, makes polycristalline films more suitable to obtain equilibrium values in the low temperature range.

We point out that this phenomenon was not appreciably observed in antimony doped specimens, very probably due to the lower diffusivity of this dopant.

The accuracy of our solid solubility data for Sb, which correspond to the equilibrium carrier density values in Fig.1, made possible to analyze in more detail the phase equilibria for the Sb-Si system.

Experimental determinations of the liquidus curve in the phase diagram were performed by

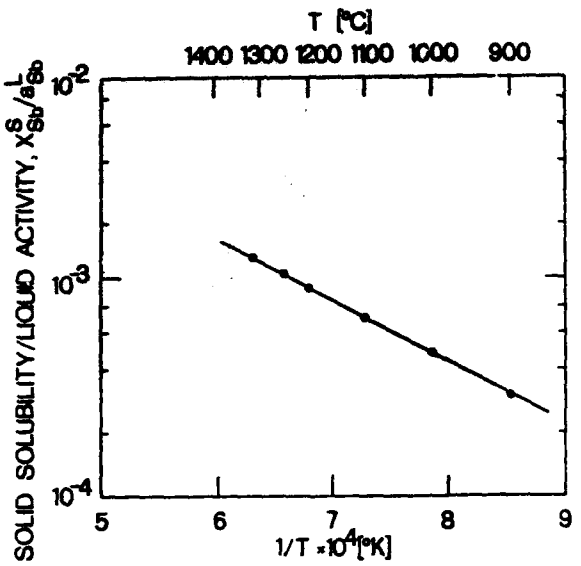


FIGURE 5

Ratio between solid solubility and liquid activity vs reciprocal temperature for antinmony in silicon.

Thurmond et al. /6/. From these data it can be deduced that a regular solution model is suitable for the liquid phase, thus allowing the calculation of the interaction parameter and hence of the activity of antimony a^L_{Sb} in the liquid /7/. The ln X^s_{Sb}/a^L_{Sb} values are reported vs reciprocal temperature in Fig.4; the precise exponential dependence which is verified in the whole temperature range is a clear confirmation of the solubility data in Fig.1. In addition this analysis provided thermodynamic data for antimony in solid solution into silicon: from the exponential dependence in Fig.5 a value of 13.4 Kcal/ mol was determined for the relative partial molar entalphy $H_{Sb}-H^o_{Sb}$; and respectively -4.7 e.u. for the relative partial molar excess entropy $(S_{Sb}-S^o_{Sb})^{xs}$.

ACKNOWLEDGEMENTS

This work was partially supported by CNR - Progetto Finalizato "Materiali e Dispositivi per l'Elettronica a Stato Solido".

REFERENCES

/1/ Nobili, D., Conf.Proc. of Satellite Symp. on "Aggregation Phenomena of Point Defects in Silicon" E.Sirtl, J.Goorissen Eds. (The Electrochem. Soc. Munich 1982) pp.189-208.

/2/ Olesinski, R.W., and Abbaschian, G.J., Bullettin of Alloy Phase Diagrams, 6 (1985) 445

/3/ Hoyt, J.L., and Gibbons, J.F., Mat.Res. Soc. Symp.Proc., "Rapid Thermal Processing", 52 (1986) 15

/4/ Olesinski, R.W., and Abbascian, G.J., Bullettin of Alloy Phase Diagrams, 6 (1985).254

/5/ Nobili, D., Carabelas, A., Celotti, G.C., Solmi, S., J. Electrochem. Soc. 130 (1983) 922

/6/ Thurmond, C.D., and Kowalchik, M., Bell Syst.Tech.J., 39 (1960) 169

/7/ Christian, J.W., "The Theory of Transformations in Metals and Alloys", Pergamon NY, Chapts. 6 and 10, 1975

DIFFUSION AND SOLUBILITY OF GOLD IMPLANTED IN SILICON

S.Coffa,L.Calcagno and S.U.Campisano

Dipartimento di Fisica,Corso Italia ,57 -Catania

G.Calleri and G.Ferla

SGS Microelettronica-Stradale Primosole - Catania

Diffusion and solubility of gold implanted in <100> p-type silicon have been investigated by Rutherford Backscattering Spectrometry and spreading resistance techniques. The gold concentration profiles are U-shaped and the concentration at the middle of the wafer thickness (C_m^s) is proportional to the square root of the diffusion time in agreement with the kick-out mechanism. The diffusion coefficient D_I^*(see text) is well described by $D_I^*=7.0*10^{-3}$ exp(-1.61/kT) cm^2/sec in the temperature range 1173-1373 K. The entropy factor associated to the ionization of the gold donor level has been determined to be 28±2.

1. INTRODUCTION

The diffusion of gold in silicon has been widely investigated [1,2] because of its technological applications such as controll of the minority carrier lifetime. The electrical parameters of Au doped silicon are of great interest in silicon power devices [3]. Gold is normally diffused in silicon starting from a thin (~300 Å) layer deposited on the surface and the concentration profile is determined by the thermal process. The introduction of gold by ion implantation will result in a better control of the gold amount in the wafer,especially close to the surfaces i.e. in the electrically active region of most devices.

2.EXPERIMENTAL

P-type <100>oriented silicon, 20 Ohm-cm resistivity is used. The wafer thickness is 620 μm and double polished wafer are used to avoid gettering of gold by the rough surface. Gold implantation is performed by means of 120 KeV Au ions and the doses are in the range 10^{12}-$5*10^{15}$ atoms/cm^2. The thermal processes are carried out under nitrogen flux in the temperature range of 1173-1373 K.

Rutherford backscattering spectrometry (RBS) of 2.0 MeV He beam is used to measure the amount of gold in the near surface region (1000 Å) of the implanted wafer,the difference between the measurements before and after the thermal process resulting in the total

amount of gold diffused into the wafer. The RBS technique cannot give informations about profiles at large depths that can be instead obtained by spreading-resistance [4] technique based on the compensating effect of the two gold levels on the silicon conductivity[5].

For spreading resistance measurements the samples are mounted on a bevel block with a bevel angle of 5°44' which gives a depth resolution of 5 μm. Two tips are then leaned on the surface with a controlled pressure and the application of a small voltage makes possible to measure the spreading resistance value.

3.RESULTS AND DISCUSSION

Spreading resistance measurements (R_s) are shown in Fig.1 for 1243 K diffused samples. The resistivity (ρ) values are obtained by appropriate calibration performed by using homogeneously doped samples and the experimental data are fitted by $R_s = 870\ \rho^{0.98}$. To convert resistivity into gold concentration it is necessary to solve the charge neutrality equation in the form

$$p + N_{Au}^{+} = n + N_{Au}^{-} + N_B \qquad (1)$$

where p and n are the holes and electrons concentration , N_{Au}^{+} and N_{Au}^{-} are the concentrations of the positively and negatively charged gold atoms and N_B is the concentration of boron in the substrate. The values of N_{Au}^{+} and N_{Au}^{-} depend on the entropy factors of both donor and acceptor (X_D ,X_A) gold related levels [6]. For the adopted wafer doping and gold concentration the gold acceptor level has a negligible influence on the final results and can be neglected.

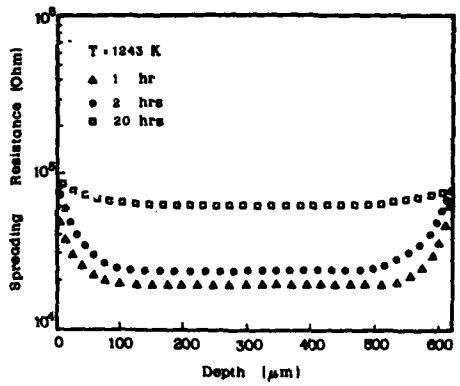


Fig.1-Resistivity profiles at 1243 K for different diffusion times.

The solution of (1) and the mobility values given in ref. 7 lead to the relation between the resistivity and the gold concentration ,which is shown in Fig.2 for different X_D values. As it appears the X_D value strongly affects the conversion resistivity - gold concentration. Because of the

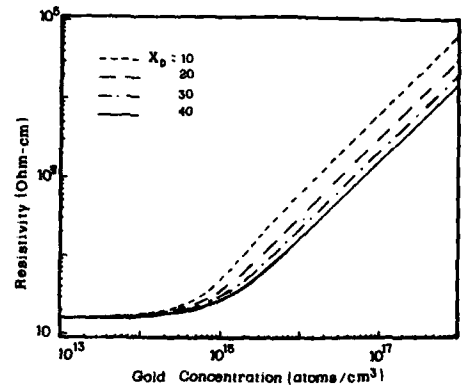


Fig.2-Silicon resistivity versus gold concentration for different X_D values.

large spread of the X_D value existing in literature [8,9,10,11] such conversion is affected by large uncertaties. Using the calculations reported in Fig.2 the resistivity profiles have been converted into concentration profiles and the total amount of gold into the wafer (area under profile) has been determined as a function of X_D as shown in fig 3. The value of the area for each thermal process is determined by RBS measurement. It is thus possible to determine the X_D value which is found to be 28±2 . Gold diffuses in silicon by migration of the fast interstitial atoms that can jump to substitutional positions. Two different mechanisms have been proposed for the interchange between interstitial and substitutional position. In the Frank-Turnbull mechanism [12] the reaction is given by

$$Au_i + V \rightleftarrows Au_s$$

where V is a vacancy. In the kick-out mechanism [13] ,

$$Au_i \rightleftarrows Au_s + I$$

where I is a silicon interstitial. The two mechanisms lead to a different trend of the gold concentration at the center of the wafer thickness (C_s^m) versus diffusion time (t). The kick-out one predicts that C_s^m increases according to

$$C_s^m = C_s^{eq} \times 2/d(\pi D_I^* t)^{1/2}$$

where d is the sample thickness , C_s^{eq} the solubility limit and D_I^* is an effective diffusion coefficient given by

$$D_I^* = D_I C_I^{eq}/C_s^{eq}$$

C_I^{eq} and D_I being the equilibrium concentration and the diffusion coefficient of silicon interstitial respectively. In Fig.4 we report C_s^m as a function of $t^{1/2}$: the agreement with the kick-out mechanism is rather good and we can estimate the effective diffusion coefficient for the investigated temperatures. An Arrhenius plot of D_I^* is reported in Fig.5 and the data are fitted by the relation

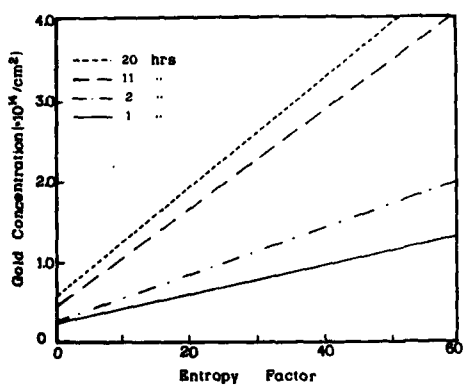


Fig.3-Total Au concentration (at./cm^2) at 1243 K as a function of X_D.

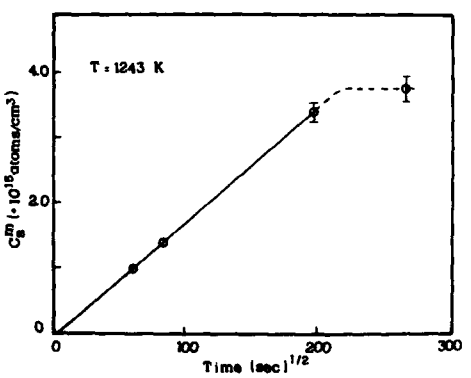


Fig.4-Gold concentration at the center of the wafer thickness (C_s^m) as a function of diffusion time for different temperatures.

$D_I^* = 7.0*10^{-3} \exp(-1.61/KT)$ cm^2/sec

shown as solid line in fig.5.

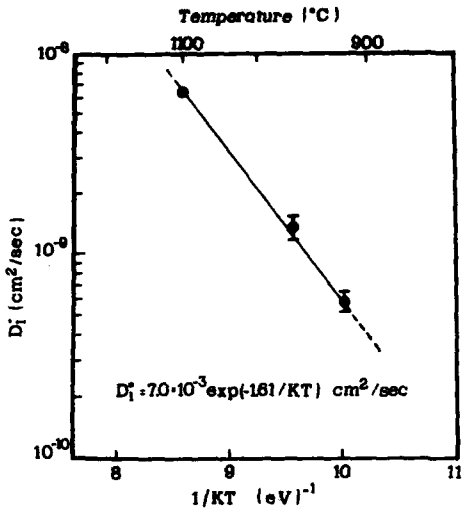


Fig.5-Arrenius plot of the effective coefficient D_I^*(see text).

4. CONCLUSIONS

By using gold implanted samples we have investigated the diffusion process of gold in silicon. Our data are in reasonable agreement with the kick-out mechanism. The activation energy for the diffusion and the entropy factor of the gold related donor level have been determined and they result 1.61 eV and 28 respectively.

ACKNOWLEDGEMENTS

The work has been supported by Progetto Finalizzato "Materiali e dispositivi per l'elettronica dello stato solido" (CNR).

REFERENCES

[1] W.M.Bullis, Solid. State El. 143 (1965) 9

[2] U.Gosele,W.Frank, A.Seeger Appl. Phys. 23 (1980) 361

[3] J.Baliga and E.Sun I.E.E.E. Trans. El. Dev. 24 (1977) 685

[4] E.E.Wang "Impurity doping processes in silicon" North Holl. Amsterdam (1981) pag. 552

[5] S.D.Brotherton, J.Bickell J. Appl. Phys. 49 (1978) 667

[6]J.A.Van Vechten and C.D.Thurmond Phys. Rev.14 (1976) 3539

[7]C.Jacoboni,C.Canali,G.Ottaviani andA.A.Quaranta Solid State El. 20 (1977) 77

[8] S.D.Brotherton,J.E.Lowther Phys. Rev.Lett. 44 (1980) 606

[9] R.Kassing, L.Cohausz, P.Van Staa, W.Mackert and H.J.Hoffman, Appl. Phys.A 34(1984)41

[10] J.A.Pals Solid. State Elec. 17 (1974) 1139

[11]N.A.Stolwijk,J.Holz and W.Frank Appl. Phys.A 39 (1986) 37

[12]F.C.Frank and D.Turnbull Phys. Rev. 104 (1956) 617

[13] U.Gosele, F.Morehead,W.Frank and A.Seeger Appl.Phys.Lett.38 (1981) 157

OPEN STENCIL MASKS FOR ION PROJECTION LITHOGRAPHY

L.-M. BUCHMANN, L. CSEPREGI, and K.P. MÜLLER

Fraunhofer-Institut für Mikrostrukturtechnik
Dillenburger Str. 53, D 1000 Berlin 33, West Germany

A processing scheme for the manufacturing of an open stencil mask has been set up by application of silicon technology and only one single X-ray lithography step for pattern generation. The mask fabrication is fully adapted to the demands of an ion projection lithography equipment by IMS. It has been proved that this mask technology permits solid structures of a complex geometry with high pattern fidelity.

1. INTRODUCTION

Ion projection lithography promises to be a successful method yielding structures in the 0.1 µm range /1/. To achieve the optimum gain from this system open stencil masks should be used which are designed to fulfill the special requirements, i.e. adapted to the ion beam divergence, the tension by thermal stress and the sputter yield by particle bombardment. A silicon membrane with a nitride top layer (Fig. 1) revealed to be most suitable for this purpose.

Considering the demands of an IMS ion projection lithography machine the fabrication of an open stencil mask will be given. Pattern generation was performed by X-ray exposition, whereas the general processing is well established in CMOS technology.

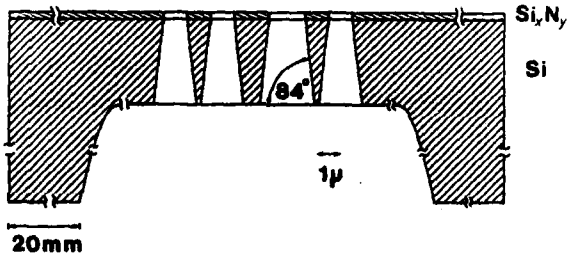


FIGURE 1

Schematic cross-section through an open stencil mask for ion projection lithography generated from a 4" silicon wafer (0.5 mm thick).

2. PROCESSING SEQUENCE

2.1 Preparation of the Membrane

The production of the thin membrane (Fig 2) followed the same process schedule as applied for the absorber masks in X-ray lithography /2/. A doted 4" silicon wafer served as the substrate on which a silicon layer of 2 t 3 µm, containing boron and germanium, was deposited epitaxially $^{+}$. Subsequently, a silico nitride layer of 0.12 µm was precipitated by LPCVD process which is essential for patter definition later on. Because the tensil strength between nitride and silicon layer i very pronounced additional "impurity" ato were implanted , which lowered the stress co siderably by rearranging the lattice of t nitride. When omitting this step, the membra cracked in the concluding etching of silic yielding debris of coiled up fragments.

The thin membrane was prepared by a tw step process. Starting with an isotropical eroding mixture of different acids (HF, HNC CH_3COOH) a smooth transfer from he origi backside of the wafer to the brim of the f was formed followed by applying an ethylene aminepyrocatechol (type S) etching agent re ving silicon till the highly boron doped la was excavated. With different covering ma windows corresponding to the geometry of intended ion transmission area could be set.

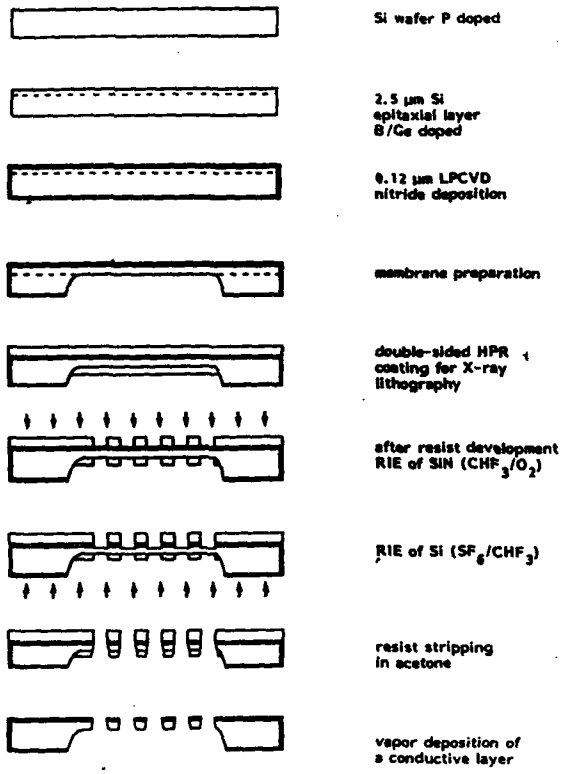


FIGURE 2

Process scheme for open stencil masks

2.2 Pattern Generation by X-Ray Lithography

The pattern was generated by exposing an HPR resist with synchrotron radiation through an absorber mask. For details concerning manufacturing of the mask and the conditions of exposure see /3/. Due to the transparency of the membrane the pattern could be transferred with coincidence into resist layers spunned on both sides of the foil by only one exposure step. The pattern fidelity of the backside test structures were routinely checked in a light-optical microscope. Fig. 3 shows a corresponding SEM micrograph.



FIGURE 3

Pattern fidelity of HPR on the backside of the membrane after exposition to synchrotron radiation and development (SEM).

2.3 Pattern Transfer - Opening of the Mask

Pattern transfer into the nitride layer of the topside of the membrane, i.e. the face the ion beam impinges on, was performed by an RIE process in a batch reactor (AME 8111). Employing a CHF_3/O_2 discharge the nitride was patterned yielding bias-free side walls without polymer depositions. Thus, the exact pattern definition could be obtained (Fig. 4).

For the purpose of mask opening a second RIE process using the fluorine containing components SF_6 and CHF_3 was determined. Etching with chlorine (BCl_3/Cl_2 or CCl_4), although generally in practice, would have required a load lock on the apparatus and, therefore, a sophisticated system to handle the fragile foils.

This process has been optimized to form the tilted sidewall corresponding to the divergence of the ion beam. The intended angle of 84° could only be obtained by sufficient sidewall passivation by species from the resist and the plasma. CHF_3 undergoes ion-molecule reactions in the plasma by which species of a high mass number (precursor to polymer formation) were originated. Therefore, the CHF_3 flow was varied in order to deposite a protecting layer, whereas the SF_6, the main source for



FIGURE 4

Pattern transfer into the nitride layer by an RIE process (CHF_3/O_2). HPR was stripped in acetone.



FIGURE 5

Sidewall protection of silicon by polymer from the resist and the CHF_3/SF_6 plasma

the highly reactive F atoms, was responsible for the amount of the etching rate. The best approach was obtained balancing these two effects with a ratio CHF_3:SF_6 of 1 (Fig. 5).

Further consideration had to be put on the thermal stress of the membrane and, especially, the resist by the etching process. Because the thermal conductivity of the delicate network is low, the transition from the wafer to the support was improved by a metal plate, to guarantee a better drain for the heat.

Due to the high selectivity of the silicon etch process with regard to the nitride an overetch could be applied to open the intended fields all over the active mask area without affecting the shape of the nitride sidewalls. Fig. 6 depicted a tapered structure obtained by etching silicon with the submitted process.

Finally, the resist was rinsed in acetone and a gold coating was deposited on the mask to prevent charging. Fig. 7 shows a 4" wafer with a 35 mm diameter membrane, which has an open area of 35%.



FIGURE 6

Tapered sidewalls of silicon test structu proposed to serve as a mask for ion project lithography. The resist has already been d solved in acetone.

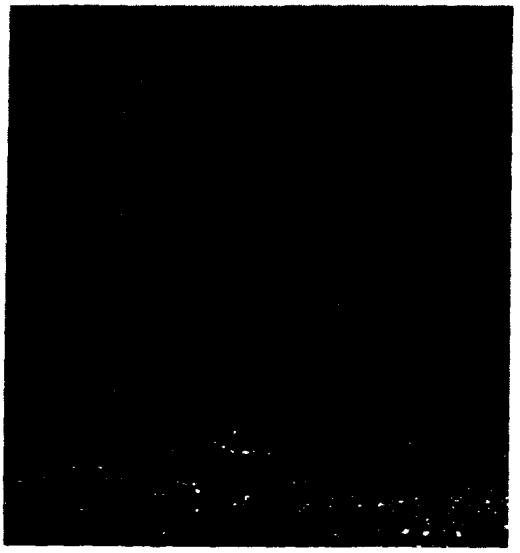


FIGURE 7

Total view on a 4" mask with an active area of 35 mm in diameter. The high transparency demonstrates the extreme thinness of the membrane.

3. CONCLUSIONS

Open stencil masks carrying test structures in the µm-range and below could be obtained by applying X-ray lithography for a self-aligned pattern generation on a membrane coated with resist on both sides and a pattern transfer splitted into two RIE processes. The exact pattern definition has been performed in a nitride layer, whereas a special profile in the supporting silicon was arranged by an etching process with compounds containing fluorine. For compensating the stress between these two layers an additional doping process of the silicon has been developed.

ACKNOWLEDGEMENTS

This work was funded by the Ministry of Research and Technology of the Federal Republic of Germany and by the Siemens AG, Munich.

NOTES

/+/ Boron in a concentration of $10^{20} cm^{-3}$ can terminate the chemical etching of silicon with an EDPS solution and thus forming the membrane. Germanium has been admixed for the purpose of internal stress compensation.

REFERENCES

/1/ Löschner, H., Ion Beam Lithography, This Conference

/2/ Hersener, J., Herzog, H.-J., Csepregi, L., Microcircuit Engineering 84, Academic Press 1985, p. 309

/3/ Heuberger, A., Microcircuit Engineering 86, North-Holland 1986, p. 3

A NEW ISOLATION PROCESS FOR VLSI DEVICES

E. Figueras*, J.L. Coppee and P. Van de Wiele

Université Catholique de Louvain, Laboratoire de Microelectronique
Place du Levant, 3
1348, Louvain-la-Neuve, Belgium

In this paper we present a new zero-bird's beak process which, with an additional photolithographic step, substitutes the thermal fully-recessed field oxide for a CVD oxide. With this process we have fabricated devices which present a very small narrow channel effect. Moreover, the use of a reference mask increases the process reproductivity and reduces the probability of the double-threshold voltage effect. The cross-section and electrical results are presented.

1.INTRODUCTION

To overcome the difficulties of LOCOS process in VLSI, several new isolation technologies [1-4], which employ a deposited oxide as a fully recessed field oxide, have been proposed. From these technologies, the BOX process [2], which uses an additionnal mask step and a double resist patterning technique prior to a planarization etch-back, is the most interesting. Nevertheless, its major drawback is that the thickness of the resist may be different for various layout designs. As a result of that, the field oxide thickness may vary in different areas [5].

In this paper, we present a new BOX process,which allows us to obtain a uniform field oxide independently of the layout design by using an additional photolithographic step and two polysilicon layers.

2.PROCESS

The starting material is <100> orientated, 14Ω.cm p-wafers. First, a pad-oxide is grown and a nitride layer is deposited. After a photolithographic step to delineate the active regions both layers are etched. The reasons to use the nitride layer will be discussed later. Then grooves are etched in the silicon substrate with a KOH solution up to a depth of 0.5 μm, using the nitride film as mask. A thermal oxidation is carried out to improve the quality o the SiO_2/Si interface. After the channel-stop implantation, a SiO_2 layer, with an equal thicknes (or little more) to that of the Si groove depth, i deposited by CVD[1] (Figure 1.a). A polysilicon laye is deposited and etched by plasma with the help o an additional resist mask which protects th polysilicon in the field areas (Figure 1.b). Th grooves between the polysilicon 1 and the oxid walls are refilled with a second polysilicon laye (Figure 1.c). Figure 1.d shows the stucture afte polysilicon 2 etch-back and SiO_2 etching. Durin the last step, the nitride, the polysilicon and th pad-oxide layers are removed to obtain th desirable structure.

Figure 2 illustrates the utility of the nitrid mask. This layer is used as a reference mask. Aft it has been removed we are sure that all the activ areas in the wafer are recovered with the san oxide thickness, independently of th misuniformity during polysilicon and oxid etching steps. The better the uniformity is, t thinner the silicon nitride layer can be. T choice of the reference mask material is limited three drawbacks: it may not be oxide, it must res KOH solution, it must be stable at high temperatu Therefore we have chosen silicon nitride.

Moreover the reference mask must assure

*B. Figueras is financially supported by the National Microelectonic Center (Education and Research Ministry, Spain)

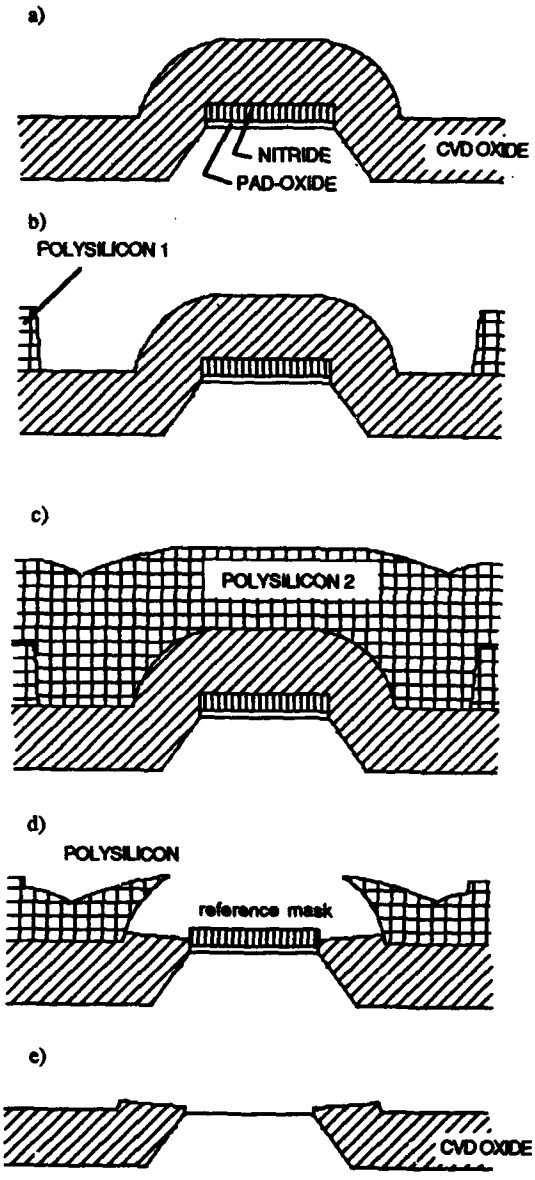


FIGURE 1

Fabrication process steps.

small positive transition step between the active and the field areas which will reduce the double threshold effect[5].

3.RESULTS

The SEM cross-sections, Figures 3 and 4, show the structures obtained following our process. It

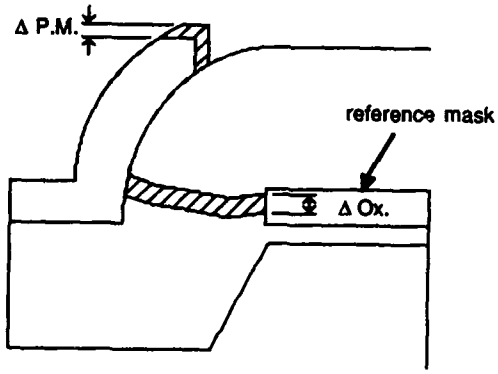


FIGURE 2

The reference mask serves to correct the misuniformity produced during polysilicon mask (ΔP.M.) and oxide (ΔOx.) etching steps.

can be seen that there's no bird's beak formation and that the field oxide thickness is uniform (except at the edge).

To present the electrical results of BOX technology, we will compare them with the results of SILO technology, wich furnish better characteristics than LOCOS[6]. Both technologies were carried out with the same test circuit and with natural transistors wich present a more important narrow-channel effect than enhanced ones due to the lower doping level of the channel.



FIGURE 3

Field oxide area. The top of the structures is covered with 0.3 μm polysilicon to evidence the oxide profile.

the isolation oxide and the poly-Si resistor was fabricated simultaneously. After the base formation, the emitter region was produced by the diffusion from the implanted N^+ poly-Si. The oxide on the base contact and P^+ poly-Si was etched away by using the photo-resist mask of emitter poly-Si etching as shown in Fig.2(C). The thick oxide was selectively grown over the heavily arsenic doped emitter poly-Si. Therfore, the oxide covering the emitter poly-Si remained after removing the thin oxide on the base contact and P^+ poly-Si, and separated the base contact from the N^+ poly-Si. The Pt-silicide was formed both on the epi-surface and P^+ poly-Si of base contact, as shown in Fig.2(D). The base electrode was fabricated with both the silicide of self-align opened base contact and the polycide, and then this can be called a salicide base contact. Opening the contact windows and Al-metallization completed the processing of SCOT transistor, as shown in Fig.2(E).

As a result of the gate speed analysis of the transistor characteristics (4), the important parameters for high speed performance are not only the cuttoff frequency f_T, the collector-base capacitance C_{TC} but also the base series resistance r_B . It was learned by the following

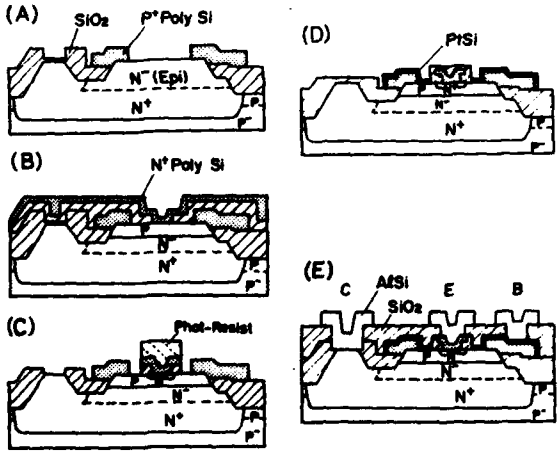


FIGURE 2

Fabrication procedure of SCOT transistor.

comparison with the coventional ISAC (8) transistor shown in Fig.1(B) that the SCOT transistor has realized the optimum transistor design with reduction of these characteristics. The r_B reduction in SCOT transistor was done by decreasing the distance D between the base contact and emitter and by the double base structure. The distance D , in the case of SCOT transistor, is determined by the salicide base contact structure and is close to 1 μm, a half value of that in ISAC one. The base area in SCOT transistor is decreased to about one-half of that in ISAC one. In order to decrease the base area, the emitter length can be decreased as the goal to maintain a small r_B. The reduction of parasitic base region, furthermore, was achieved by the full walled base structure with the semi-recessed oxide and by the salicide base contact structure. The vertical down-scaling, in which the emitter depth is 0.1 μm and the base width is 0.13 μm, produces the higher f_T. Table 1 shows that a SCOT transistor has been synthesized in the transistor design for high performance.

3. GATE SPEED AND PRESCALER IC

The performance of SCOT transistor used for a prescaler IC was improved as compared with the current product with ISAC process, that is, the C_{TC} and r_B decreased to half value and f_T became twice as high, as shown in Table 1. The maximum f_T obtained, moreover, is 9.5 GHz at I_C = 6 mA by the SCOT transistor in which the emitter is

TABLE 1

Comparison of features of newly developed prescaler and current prescaler.

		New development	Current product
Process technology		SCOT	ISAC
Emitter size		1.5 x 3 μm^2	1.5 x 5 μm^2
Emitter depth		0.1 μm	0.4 μm
Capacitance	C_{TC}	9 fF	20 fF
Base resistance	r_B	49 Ω	92 Ω
Cutoff frequency	f_T	4.1 GHz	2.2 GHz
Delay time of Ring-osc.		140 ps	267 ps
Max. operating frequency		2.1 GHz	1.1 GHz
Power dissipation		56 mW	125 mW

four-fingers of 1.5 x 5 μm^2.

It is demonstrated in Fig.3 that the gate speed t_{pd} of the ECL ring-oscillator employing SCOT transistor on same emitter size (1.5 x 5 μm^2) was faster than that of ISAC one, especially at the high current range by effect of reducing r_B. The mimimum t_{pd} achieved 116 ps at gate current I_g = 1.3 mA. Simulating t_{pd} in scaling down of the SCOT transistor to 1.0 μm emitter, furthermore, the high speed can been realized at less than 80 ps.

A 1/128,1/129 two-modulus prescaler IC was fabricated with 1.5 μm SCOT transistors and poly-Si resistors. First level metallization of AlSi and polyside cross-under interconnection were employed. This prescaler IC operated in a wide range from 400 MHz to 2.1 GHz with 56 mW at 5-V supply voltage. In Fig.4, the performance of SCOT prescaler IC was compared with other products. The SCOT prescaler IC had four times higher performance than the ISAC prescaler. This prescaler IC operated with the half power dissipation of the reported Si prescaler (1) and with the same one of the GaAs prescaler (2). Decreasing the power dissipation, we obtained 1.4 GHZ with 30 mW and 850 MHz with only 19 mW.

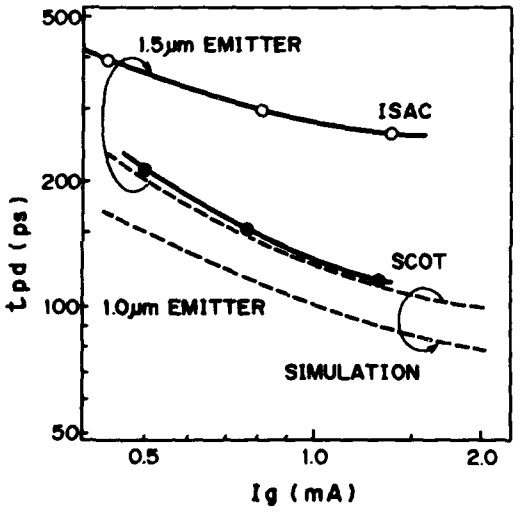


FIGURE 3

Relations between gate speed and gate current.

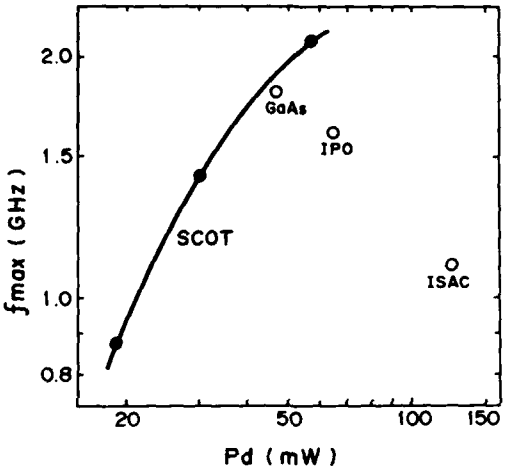


FIGURE 4

Comparison of maximum operating frequency and power dissipation.

4. VSC MASTERSLICE

The actual pattern layout of two-input OR/NOR gate is shown in Fig.5. In the case of the current cell approach, a relatively large number of the nonutilized elements ramain, for example the macrocell array MCA in the simple logic functions. The VSC concept is based on the design of an array which is constructed from cellular units. This VSC construction was found to reduce the nonutilized elements.

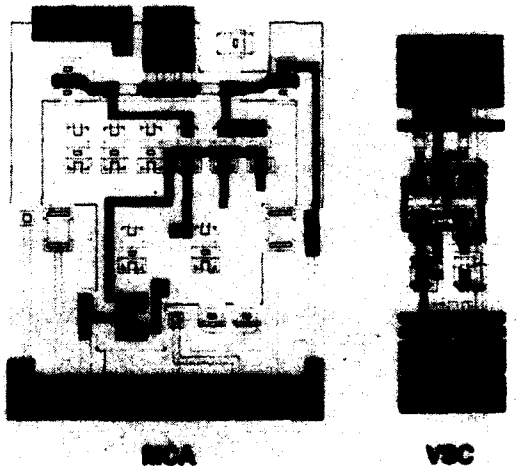


FIGURE 5

Layout pattern of two-input OR/NOR gate.

Fig.6 shows a schematic diagram of SCOT transistor and polycide interconnection used for VSC masterslice. To realize a VSC structure, the proper use of poly-Si patterns was required during the slice process. The resistor value in each logic cell was determined by the silicidation of poly-Si pattern. In addition, unused poly-Si patterns can be utilized for the policide interconnection.

An ECL 18K-gate masterslice (7) was developed by VSC approach and fabricated by employing 1.5 μm SCOT process with four-level metallization. The features of the VSC masterslice are summarized in Table 2. The gate density was increased by more than 20 % in the VSC structure compared with the current cell structures. The basic gate delay of 150 ps was attained at the power dissipation of 2.4 mW.

5. CONCLUSION

As the excellent structure with reduction of C_{TC}, f_T and r_B simultaneously, the SCOT process by employing self-alignment silicide technology has been proposed. A two-modulus prescaler IC has achieved 2.1-GHz operation with 56-mW power dissipation. An ECL 18K-gate masterslice has been developed by a VSC approch which maximized the utilization of elements.

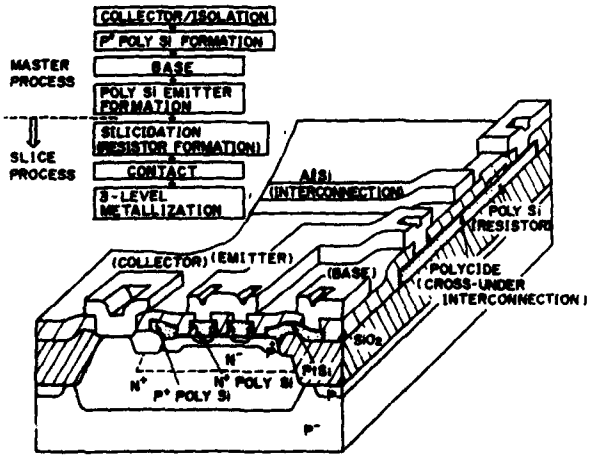


FIGURE 6

Schematic diagram of SCOT transistor and polycide interconnection used for VSC masterslice.

TABLE 2

Features of VSC masterslice.

Technology	1.5 μm rule SCOT
No. of transistors	39,936
No. of poly-Si resistors	53,248
No. of units	13,312
Unit size	24 μm X 204 μm
Metal pitch	1st 8 μm
	2nd 6 μm
	3rd 8 μm
No. of channels	1,936
No. of I/O pins	256
Interface	ECL 100K compatible
Intrinsic gate delay	150 ps
Supply voltage	V_{EE} : -4.5 V
	V_{TT} : -2.0 V
Switching current	0.4 mA
Emitter-follower current	0.3 mA/0.6 mA
Chip size	11.90 mm X 11.96 mm

ACKNOWLEGEMENTS

The author would like to thank N.Katoh and T.Nishimura for circuit deign, and K.Sakaue and Y.Kinosita for wafer processing. We would also like to thank N.Tsubochi for fruitful discussions and express gratitude to K.Shibayama and H.Nakata for their encouragement.

REFERENCES

(1) H.Suzuki, T.Akiyama and K.Ueno, IEDM Tech. Digest (1984) 682

(2) K.Maemura, T.Takahashi, S.Inoue, Y.Mitsui, S.Orisaka, O.Ishihara and M.Otsubo, IEDM Tech. Digest (1985) 94

(3) C.Y.Ting, IEDM Tech. Digest (1984) 110

(4) T.Hirao, T.Ikeda and N.Katho, Extended Abstr. of 17th Conf. SSDM (1985) 381

(5) W.Brackelmann, H.Fritzsche, H.Ullrich and A. Wieder, IEEE J. Solid-State Circuits, SC-20 (1985) 1032

(6) M.Tatsuki, S.Kato, M.Okabe, H.Yakushiji and Y.Kuramistu, IEEE J. Solid-State Circuits, SC-21 (1986) 234

(7) T.Nishimura, H.Sato, M.Tatsuki, T.Hirao and Y.Kuramitsu, IEEE J. Solid-State Circuits, SC-21 (1986) 727

(8)Y.Akasaka, Y.Tsukamoto,T.Sakurai, T.Hirao, Y.Horiba, K.Kijima and H.Nakata, IEDM Tech. Digest (1978) 189

TRENDS IN HETEROJUNCTION SILICON BIPOLAR TRANSISTORS

R. MERTENS, J. NIJS, J. SYMONS, K. BAERT and M. GHANNAM

Interuniversity Microelectronics Center (IMEC)
Kapeldreef 75
B-3030 Leuven, Belgium

Different types of bipolar transistor emitters are described. Epitaxial emitters can be achieved by solid phase epitaxial regrowth of polysilicon (at T > 850°C) and recently by glow discharge deposition at T = 250°C and recrystallization (at T = 700°C). Wide band gap emitters and narrow bandgap bases result in very high emitter efficiency which has to be traded-off with emitter and base series resistances.

1. INTRODUCTION

VLSI bipolar transistors typically have polysilicon emitters to achieve high emitter efficiency and large packing density. In recent years [1,2] it has become clear that for optimum performance of polysilicon emitters in VLSI applications a trade-off must be made between emitter efficiency and emitter series resistance. Work at several laboratories has indeed clearly shown that the emitter Gummel number (GN_E) is not the appropriate figure of merit of a modern emitter-base junction used in high speed VLSI applications. The true figure of merit depends on GN_E and also on the emitter series resistance.

In this paper different emitter-base structures aiming at high emitter efficiency and low emitter series resistance will be discussed. In a first part of the paper our work on epitaxial emitters will be described. Although epitaxial emitters are not heterojunctions in the strict sense of the word, they are included here because epitaxy is a technology closely related to heterojunction processing and since in the short term these emitter may turn out to be the best alternative for poly-emitters.

The second part of the paper will deal with true heterotype emitters starting with an overview of the different wide-gap emitters that have been proposed in the literature. Finally, the paper will end with a discussion about the possibilities of narrow base transistors.

2. SILICON BIPOLAR TRANSISTORS WITH EPITAXIAL EMITTERS

As pointed out in the introduction polysilicon emitters are suffering from a trade-off that has to be made between emitter efficiency and emitter series resistance. Recent work [1] has indicated that high emitter efficiency only can be realized at the expense of a large series resistance. Such a specific series resistance R_E (ohm.cm^2) is detrimental for high speed performance if [3]:

$$R_E > \frac{kT}{J} \qquad (1)$$

In (1) J is the emitter current density which in advanced bipolar transistors can be as large as 10^5 A/cm^2. Formula (1) predicts that R_E should be lower than 2.6×10^{-7} A/cm^2 to eliminate degradation of the transconductance. This is a very low value which is difficult to achieve with polysilicon emitters. The emitter series resistance of a polysilicon emitter is caused by two components : an interface resistance caused by the presence of a thin interfacial oxide layer between the poly and mono-silicon and a true contact

resistance at the metal-poly interface. Both components are considerably higher than in the case of mono-crystalline silicon emitters. The first one does not exist in a monocrystalline emitter and the second one strongly depends on the surface free carrier concentration. This latter is considerably smaller in the case of a poly-emitter due to carrier trapping and impurity atom segregation at the grain boundaries.

Besides series resistance problems, the thickness of the interfacial oxide layer is not easily controlable which causes serious yield and reproducibility problems in poly-emitter structures. The elimination of this thin layer is therefore two fold advantageous at the expense of reduced current gain.

The reduction of emitter series resistance, coupled with an expected increase in reproducibility is the prime reason for the interest in epitaxial emitter structures. In addition, due to the fact that epitaxial emitters are deposited doped with the appropriate impurity and not doped by compensation, a better emitter efficiency is expected.

2.1. Epitaxial Regrowth of Polysilicon

It has been demonstrated that it is possible to align polysilicon epitaxially to the underlying single crystalline substrate by high temperature furnace annealing or rapid thermal annealing [4-7]. In order to start the alignment process, a direct contact between the polysilicon layer and the single crystalline substrate must, however, take place. In other words the native interfacial oxide, though very thin, should be removed. Figure 1 displays a cross sectional TEM photograph of a partially aligned polysilicon film after a 30' anneal in Argon at 1000°C, and shows clearly how discontinuous the interfacial native oxide layer becomes. This readily occurs at annealing temperatures above 950°C, and can be stimulated by annealing in an oxidizing ambient if the anneal is to be performed at lover temperatures (~ 850°C) [4]. Proper adjustment of the annealing time, temperature and ambient can lead to an epitaxial regrowth of the entire polysilicon film. Undoped, n-type and p-type polysilicon films have shown similar behavior [4-6] which makes low resistivity epitaxial burried p^+ base contacts for self aligned "poly --> epi" bipolar transistors possible.

It is not possible, however, to obtain emitter box-shaped profiles using this method because of impurity redistribution in the aligned region as well as impurity out-diffusion into the substrate during annealing.

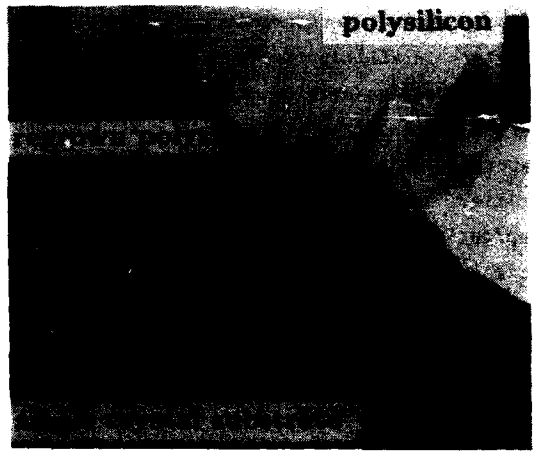


FIGURE 1

Cross sectional TEM photograph showing epitaxial alignment of polysilicon to the underlying substrate after anneal in Argon for 30' at 1000°C (1cm = 80 nm).

2.2. Low Temperature Epitaxial Grovth by Glow Discharge Deposition

Another very attractive approach to form a low temperature epitaxial emitter has recently been proposed by the authors [8]. Under proper pre-deposition cleaning conditions, heavily doped amorphous silicon films deposited from a silane plasma at 250°C on a single crystalline silicon substrate will

epitaxially recrystallize throughout the film after an annealing at a temperature of 600-700°C during typically 30'. Figure 2 shows a cross-sectional TEM picture of the as-deposited amorphous silicon film without subsequent heat treatments. This picture clearly shows that over large fractions of the interface the silicon is initially deposited under single crystalline form epitaxially aligned with the substrate. The thickness of this layer depends from point to point and is typically 10 nm. The remaining part of the layer is deposited in the amorphous state. The existance of this thin epitaxially aligned layer is probably due to a combined effect of the presence of F^- ions at the surface resulting from the pre-deposition cleaning step, and the reducing effect of the silane plasma on the "native" oxide during growth. Figure 3 shows the same example after annealing at 600-700°C. Clearly the entire n-type layer has recrystallized up to the surface; the presence of twin defects can be observed. The advantages of this process are the large throughput, the good homogeneity over large areas, the effectiveness of dopant introduction and the low processing temperature. Transistor operation with very reasonable efficiency (emitter Gummel number = 10^{14} $cm^{-4}s$) has been obtained (Fig. 4).

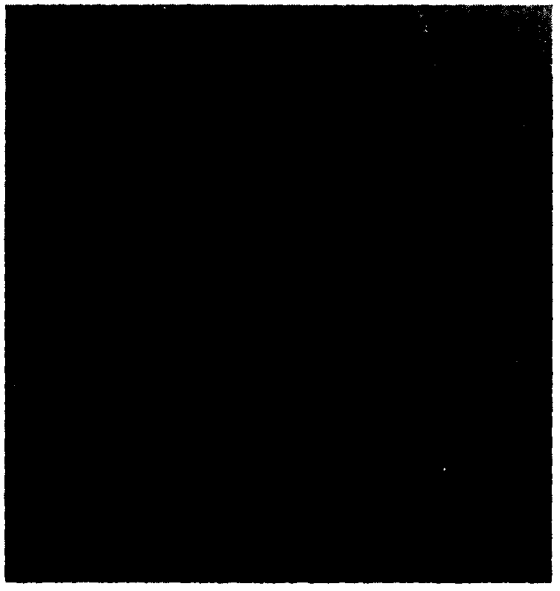


FIGURE 2

HRTEM photograph showing the as-deposited n^+ a-Si:H/c-Si interface. The deposited Si is epitaxially aligned to the substrate over large fractions of the interface.

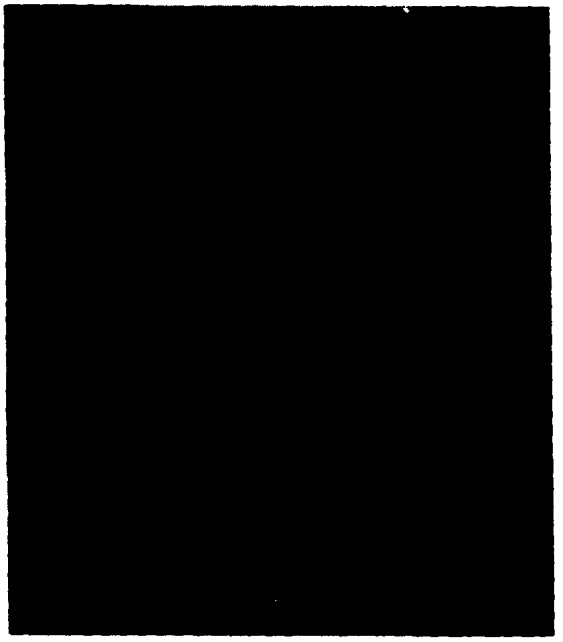


FIGURE 3

HRTEM photograph showing the n^+ a-Si:H/c-Si interface after 30' anneal at 700°C. The entire n^+ a-Si:H film is epitaxially aligned to the substrate. The insert represents the microdiffraction pattern of the recrystallized region.

3. WIDE BANDGAP EMITTER SILICON BIPOLAR TRANSISTORS

Wide bandgap silicon bipolar transistors are heterotype silicon bipolar devices which may allow ultra high f_T ($\geq$ 40 GHz) and room temperature gate delays of 5 ps [9]. These expectations are based upon the following advantages : the combination of a sufficiently large current gain with a low intrinsic base

resistance, the elimination of emitter stored charge and a better control on the uniformity of ultra-shallow box-type emitter base profiles. As discussed earlier these advantages should not be realized at the expense of a large emitter resistance, causing a decrease in transconductance and completely annihilating the expected gain in transistor performance.

At the moment most of the work on silicon hetero-type devices is still in the early development stage since only dc results are largely available; results on the dynamic performance of silicon heterojunction devices arevery limited.

The challenges in wide-gap emitter research can be summarized as follows :

- to find a wide-gap material compatible with state-of-the art silicon processing
- the doping efficiency of the wide-gap material must be sufficiently large such that low bulk resistivities and contact resistances can be obtained.

Unfortunately it turns out that these two requirements are not easily compatible. The following materials have been investigated, with variable success, as wide gap emitter :

- GaP : this material is lattice matched to Si but suffers from interface doping effects; so far poor transistor performance has been reported [10].
- SIPOS : yields excellent GN_E but large R_E. Experiments have shown that current gain is not related to the wide emitter gap but to the presence of a thin interfacial oxide layer [11].
- β-SiC : although not lattice-matched on silicon good quality epitaxial growth without large built-in stress has been demonstrated [12]. This material is very promising, although device performance has not yet been reported.
- Amorphous Si (a-Si:H) : excellent GN_E but large R_E (> 10^{-3} Ωcm^2) [13]; large series resistance is caused by contact rather than by the presence of an interfacial oxide layer. The TEM picture of Fig. 2 demonstrates the onset of epitaxial alignment over a large fraction of the interface at the growth temperature of 250°C in the plasma CVD reactor. This indicates that the large emitter Gummel number is not an interface but a true heterojunction effect. Amorphous SiC:H has also been tried [14], but devices suffer from similar problems as with a-Si:H.
- Microcrystalline silicon [15] : allows large bandgap (1.4 eV) and reasonably low resistivities (10^{-2} Ωcm) to obtain low series resistance. However, growing this material reproducibly using the conventional plasma CVD without epitaxial alignment may be difficult.

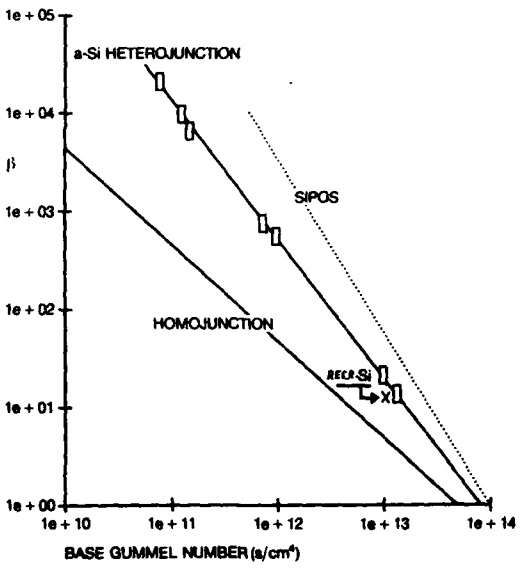


FIGURE 4

Max DC current gain versus base Gummel number for different types of emitters.

Figure 4 presents some experimental results indicating an increase in the dc current gain β when Si wide bandgap emitters are used. From these preliminary results it follows that, in the opinion of the authors, two

"wide-gap" materials should be considered as serious candidates : SiC and to a smaller extent microcrystalline silicon.

4. NARROW BASE SILICON BIPOLAR TRANSISTORS

It is now well known [16] that the Ge_xSi_{1-x} layers can be grown pseudomorphically on silicon if the layers are thinner than the critical thickness. If such a p-type layer is overgrown by an n-type doped Si layer a narrow base transistor is formed. Initial calculations indicate that with x = 0.2 a large bandgap difference between emitter and base and a critical thickness large enough to avoid punchthrough and yield low base resistance, can be combined. As in this structure the silicon in the emitter will remain cubic and unstressed the bandgap difference will be almost completely seen as a bandoffset at the valence band edge yielding a very efficient np heterojunction.

In the opinion of the authors such a narrow-gap base heterotransistor is extremely promising for the following reasons :

- emitter and collector interchangeability
- smaller turn-on voltage, allowing a somewhat lower power dissipation
- due to the absence of collector stored charge a high speed saturated logic can be developed
- no problems with low emitter specific contact resistance.

5. CONCLUSIONS

Directly diffused or implanted emitters are not compatible with very thin base VLSI bipolar transistors. Polysilicon emitters suffer from high emitter resistance and in the short term epitaxial emitters seem to be the best alternative. On the other hand, more research sould be oriented towards narrow bandgap base heterojunction bipolar transistors as they present comparably high emitter efficiency to wide band gap emitters with, however, no emitter resistance problems.

6. ACKNOWLEDGEMENT

The authors would like to thank J. Vanhellemont for the TEM analysis.

REFERENCES

[1] Crabbé, E., Swirhun, S., del Alamo, J., Pease, R.F. and Svanson, R.M., IEDM Tech. Dig. (1986) 28.

[2] Stork, J.M.C. and Cressler, J.D., Symp. on VLSI Technol. Dig. Tech. Papers (1986) 47.

[3] Solomon, P.M., Proceedings of the IEEE 70 (1982) 489.

[4] Ghannam, M.Y. and Dutton, R.W., Proceedings of the IEEE Bipolar Circuits and Technology meeting (1986) 5.

[5] Tsaur, B.Y. and Hung, L.S., Appl. Phys. Lett. 37 (1980) 648.

[6] Wong, C.Y., Michel, A.E., Isaac, R.D., Kastl, R.M. and Mader, S.R., J. Appl. Phys. 55 (1984) 1131.

[7] Natsuaki, N., Tamura, M., Miyazaki, T. and Yanagi, Y., IEDM Tech. Dig. (1983) 662.

[8] Baert, K., Symons, J., Vandervorst, W., Vanhellemont, J., Caymax, M., Poortmans, J., Nijs, J. and Mertens, R., to be submitted to APL for publication.

[9] Wieder, A.W., IEDM Tech. Dig. (1986) 8.

[10] Kroemer, H., Proceedings of the IEEE 70 (1982) 13.

[11] Oh-uchi, N., Hayashi, H., Yamoto, H. and Matsushita, T., IEDM Tech. Dig. (1979) 522.

[12] Yamanaka, M., Daimon, H., Sakuma, E., Misawa, S. and Yoshida, S., J. Appl. Phys. 61 (1987) 599.

[13] Ghannam, M., Nijs, J., Mertens, R. and De Keersmaecker, R., IEDM Tech. Dig. (1984) 746.

[14] Sasaki, K., Furukava, S. and Rahman, M., IEDM Tech. Dig. (1985) 294.

[15] Ghannam, M., Nijs, J., De Keersmaecker, R. and Mertens, R., EPS abstracts, ESSDERC (1985) 46.

[16] Jorke, H. and Herzog, H.-J., J. Electrochem. Soc. 133 (1986) 998.

Session B2.2

SOI Workshop II

Chairman: D. Mc Caughan

Tuesday, September 15, 1987

ELECTRICAL PARAMETERS OF SOI MATERIAL OBTAINED BY ZMR AND OXIDIZED POROUS SILICON

M. Haond, G. Bomchil, J-L. Regolini, D. Bensahel, D. Dutartre, D-P. Vu, K. Barla, H. Halimaoui, R. Herino.

Centre National d'Etudes des Telecommunications, BP 98, 38243 Meylan Cedex, France

A. Monroy, S. Thouret, Y. Gris.

Thomson Semiconducteurs, BP 217, 38019 Grenoble Cedex, France

Lamp-Zone Melting Recrystallization (ZMR) of deposited silicon on oxide has proved to be suitable for making devices. We present here electrical results obtained in this material on batches of 4-in. wafers, which confirm its crystalline quality. We also present recent results obtained in SOI material prepared by oxidizing a buried porous layer. Since laser-ZMR is still in the race toward the fabrication of a material compatible with 3-D circuits fabrication, some new results are periodically available. By presenting electrical results of the three types of material, we compare and discuss the future trends in SOI concerning each of these three techniques.

1. INTRODUCTION

Silicon-On-Insulator (SOI) is expected to help solving some of the present problems encountered in the CMOS technology of VLSI circuits, namely in the race toward smaller dimensions: latch-up, dielectric isolation and density of integration, power consumption,... Many techniques have been investigated. The use of an oxygen ion implantation (SIMOX) is presently widely studied over the laboratories, and is discussed elsewhere in these proceedings. We have been working on Zone Melting Recrystallization (ZMR) for a long time /1/. This technique has therefore reached a level where batches of 4-in. wafers can be processed reproducibly and provide a device-worthy material /2/. Another technique is attractive since the processing apparatus is simple and cheap: the oxidation of porous silicon obtained by anodizing a N/N+/N structure. The last technique we report below is laser-ZMR which is one of the techniques which should be used for making 3-D devices. We present these three techniques and compare their respective electrical parameters as measured after having ran wafers of each type in a 3µm CMOS technological process.

2. SOI MATERIAL PREPARATION

The preparation of lamp-ZMR SOI material has already been reported /3/. Shortly, a grating is etched in an oxidized wafer. It consists in 0.4 µm deep and 36 µm wide stripes separated by 4 µm lines. A 0.5 µm to 0.6 µm thick encapsulated polysilicon film is used. The cap is a 1.6 µm thick oxide. It avoids the delamination of the liquid silicon and limits mass transport at melting. The lamp apparatus /3/ consists in a row of halogen lamps used to preheat the wafer up to 1150 °C. An additional lamp placed in an elliptical mirror is used to melt the deposited polysilicon film. The molten line is scanned at a speed of 0.2 mm/s. The grating etched in the underlying oxide provides an effective defect entrainment, thanks to a solidification front modulation /4/. Fig.1a presents a SEM micrograph of a cross-section of a typical lamp-ZMR wafer. Extended characterizations of these samples have already been reported /5/. The remaining defects such as grain-,subgrain boundaries (GBs, SGBs) or precipitates are entrained upon the 4 µm steps of the relief grating. The wafers are flat and the slip lines if there are any are located in the bulk substrate and not in the SOI film itself. Batches have been processed which

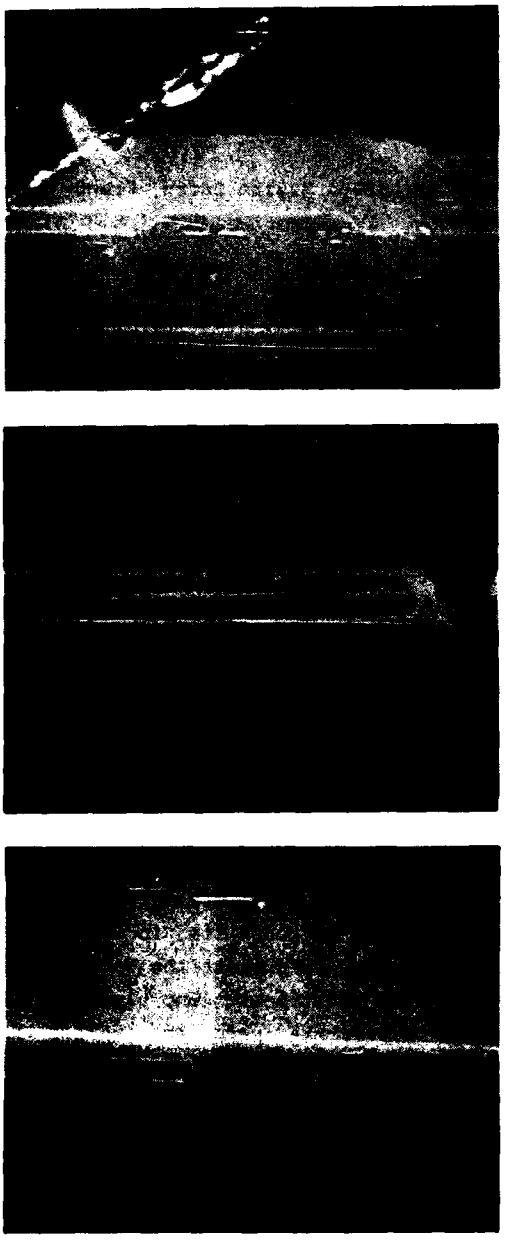


FIGURE 1

SEM cross-sectional micrographs of:
a) a lamp-ZMR oxide capped sample near a defect entrainment line
b) an oxidized porous sample near an anodization access window . The sample has been etched in a 1:7 HF:FNH4 solution
c) a laser-ZMR sample near a SEG-filled seed. The sample has been immersed in a buffered HF solution.

show the reproducibility of the lamp-ZMR procedure as well as the compatibility of the wafers with a CMOS process line.

A different approach has been studied to fabricate SOI material: it is based on the transformation of single crystal silicon into porous silicon by anodization in an HF electrolyte. By taking advantage of the preferential anodization of N+ silicon, a buried porous layer can be obtained by opening windows in the N-type epitaxial layer of an N/N+/N structure /6/. The formation of porous silicon proceeds laterally until the whole volume of the buried layer is transformed, in a self-limited reaction. We have been working with a controlled potential in a potentiostatic three electrodes configuration. The anodization conditions in HF electrolytes solutions are adjusted to give an homogeneous porous layer of about 56% porosity. An oxidation step provides a buried oxide which is equivalent to a thermal oxide /7/, regarding its resistance to chemical etching in HF: FNH4 solutions. Fig.1b is a SEM micrograph of a cross-section of the SOI strucure after the porous layer has been oxidized. The buried oxide is homogeneous, displays flat and abrupt interfaces. In this sample, the N+ layer doping level was 1.5 E19 cm-3 (antimony) and was anodized in 35% HF solutions in ethanol. The oxidation procedure has been described elsewhere /7/. By using this procedure, 4-in. SOI wafers have been fabricated which are compatible with a standard CMOS technological process. The SOI material consists in 4/36 µm window/active region stripes, all across the wafer.

Laser-ZMR has been a pionneering technique for the obtention of SOI material, but its application has been restricted because of several drawbacks. The spot size and related overlapping problems remain. Since the two previous techniques cannot afford the opportunity of making 3-D devices, the laser has not been abandonned. Therefore, new efforts have been made on the way toward obtaining large defect-free areas. In

laser-ZMR where a seed is used, defects are related to the thermodynamical behaviour of the seed area upon melting, i.e. the volumetric contraction of silicon and the temperature gradient between the seed and the SOI region. These effects can be overcome by using a seed filled with single crystal Si or by using discontinuous seeds /8/.

In our case, we have used a Selective Epitaxial Growth process (SEG) /9/. Our samples are as follows: 4-in. wafers are oxidized up to 0.5-0.6 µm and then etched by Reactive Ion Etching (RIE) for delineating the seed regions: 2 or 4 µm wide and 40 µm pitch. The seed lines are <100> or <110> oriented. The solidification front is here controlled by the trailing edge of the molten spot.

A solidification front parallel to the <110> direction is therefore obtained by using an elliptical spot slanted at 30 ° from the <100> scan direction.

A 1 µm poly-Si film is then deposited, followed by the deposition of a 1 µm oxide cap. Since the slanted elliptical spot is scanned parallel to the seed lines, the defects are rejected near one end of the seed. Fig.1c shows a SEM micrograph of a cross-section of our sample near the seed region. The high crystal quality has been confirmed by TEM. The crystalline defects are the SGBs which correspond to about 1 degree of misorientation between each side of the SGB. Similar results have been obtained using discontinuous seed /8/. Isolated dislocations exist in the seed regions. They do not extend very deep into the substrate and their density is lower with 2 µm seeds than with 4 µm openings.

3. ELECTRICAL PARAMETERS

The three types of SOI wafers have been processed in similar self-aligned 3 µm poly-Si gate CMOS sequences. The main technological parameters are summarized in Table 1. The laser results presented below have been obtained on samples prepared and recrystallized by the LETI and using discontinuous seeds /8/. Natural transistors are

SOI	Gate Oxide	Lateral Isolation	Gate Material
5000 A	420 A	RIE of mesas	Poly-Si:4200 A

Table 1 : Process Parameters.

fabricated by avoiding the channel implants, i.e. using the as-prepared SOI material. They can either be enhancement or depletion mode transistors. These natural transistors provide an adequate tool to study the properties of the interfaces and to measure the residual doping concentration, by C-V, I-V and transconductance (gm) measurements /10/. Since the SOI film thickness is about 5000 A, the channel will be completely depleted or not, depending on the residual doping level of the SOI. Table 2 presents values of drain current as measured on depletion N+/N/N+, W/L= 30/30µm (natural) transistors at Vg= -2V and Vd= 5V. It is below E-13 A/µm in the lamp-samples whereas it is about E-9 A/µm in the laser-samples and E-6 A/µm in the FIPOS samples. These values are attributed to the residual doping level and to interface states at the back interface. In the three types of SOI material, the residual doping is N-type. It is in the low E15 cm-3 in the lamp samples and in the high E15 cm-3 in the laser samples. No satisfactory explanation has been found to explain this N-type, whereas it is in the E16 cm-3 for the FIPOS samples where it can be attributed to a diffusion or to any incorporation mechanism of species from the N+ layer used for the preparation of the material (N/N+/N structure). In order to better qualify the material quality, we have derived the minority carrier lifetime from the time to form the inversion layer in a depletion mode transistor /11/. Table 2 summarizes the orders of magnitude obtained in each type of SOI. It shows that the lamp material is approaching bulk material quality whereas FIPOS and laser materials have traps which reduce the lifetime. As shown in Table 2, we have also derived values of electron and hole mobilities from gm(Vg) curves at Vg=

	Residual Doping	tg	μe	μp
Lamps	N : 2-3 E15 cm-3	10 μs	900	230
FIPOS	N : 2-3 E16 cm-3	1 μs	500	145
Laser	N : 7-8 E15 cm-3	1 μs	1000	200

tg: Minority Carrier Lifetime.
μe, μp: Electron, Hole Mobility.

Table 2 : Electrical Parameters Obtained from Natural Transistors (N+/N/N+ and P+/N/P+).

Vfb (flatband voltage) /12/. Notice that these values do not correspond to those measured in inversion or accumulation layers. They are close to the bulk values /13/ whereas the second ones are lower as a result of increased scattering.

These results are confirmed by measurements performed on enhancement mode transistors (where the channel has been ion implanted). In Table 3, we present leakage current levels as measured on 72/5 μm N-type edgeless transistors at Vg= -2V and Vd= 5V. In each type of SOI, they remain below 1 pA/μm of channel width. Table 3 also displays surface mobility values as derived from Id-Vg curves in the linear region. The high value of the electron mobility together with a low hole mobility in the laser-samples could be due to some residual local stress within the SOI film. The low hole mobility of the FIPOS sample could result from a poorer quality either of the interfaces or of the material as compared to the other materials. These parameters have been derived from measurements performed on edgeless transistors since these allow to avoid the parasitic edge channel encountered in mesa etched N-type transistors /14/. Fig.2 shows Id-Vg curves of both P- and N-type edgeless 72/5 μm transistors obtained in lamp-ZMR wafers at Vd= 0.1 V. The subthreshold slope is about 120 mV/decade and confirms a low density of interface states. This slope can overcome the theoretical "ideal" value as a result of the "kink" effect /15, 2/due to the floating substrate /16/. It can be avoided by connecting the source and the channel or reduced by thinning the SOI film /17/.

249-stage, 3.5 μm effective gate length ring oscillators have been fabricated, in order to test the dynamic characteristics of the materials. Table 3 summarizes the results for the three SOI techniques. We believe that in the laser-samples where the mobility on individual transistors is high (see Table 3), the increased propagation delay time is partly due to the thin oxide used to isolate the seeds (gate oxide), and also to the low thickness of the underlying oxide (0.3 μm).

	Leakage Current (•)	μe cm2/V.s	μp cm2/V.s	tp ns
Lamps	<pA	620	220	1
FIPOS	<pA	570	155	1.7
Laser	<pA	570	185	2.2

(•) Measured at Vg= [illegible] and Vd=+/-5V.
μe, μp: Electron, Hole Mobilities.
tp: Propagation Delay Time per Gate (3.5 μm effective channel length).

Table 3 : Electrical Parameters Obtained from N- and P-Enhancement Mode Transistors .

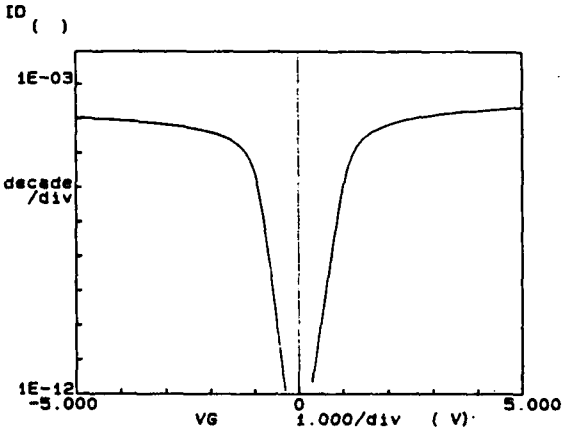


FIGURE 2

Id-Vg curves of P- and N-channel transistors fabricated in lamp-ZMR material. Channel width, W= 72um and gate length, L= 4um. Vd= +/- 0.1V and Vg= -/- 5V. The Bulk substrate voltage Vb is 0V.

The number of processed wafers is too low for the laser- and FIPOS-SOI material to provide reliable statistical studies from a wafer to another or from a run to another. It must be pointed out, however, that in single wafers the laser- and FIPOS-SOI film thickness is very uniform. This point is very important for technological steps such as SOI film thinning, dielectric isolation (mesa etching or LOCOS definition), energy adjustment of the Drain/Source implants...

Statistical results are available for the lamp-SOI material where many batches have been processed. They have first shown that the residual doping level should be reduced down to the £15 cm-3 or less in order to allow the adjustment of the threshold voltage of both N- and P-type transistors. Notice in Fig.2 the symmetrical characteristics of both P- and N-channel transistors. In Fig.3, we display the threshold voltage distribution of enhancement mode 20/3 µm N-type transistors obtained on 4 4-in. wafers of a same batch. The mean value is 0.9V with a standard deviation of 61 mV. We have already presented some statistical results, namely on the influence of the defect entrainment lines and shown how they scatter the threshold voltage distribution /2/.

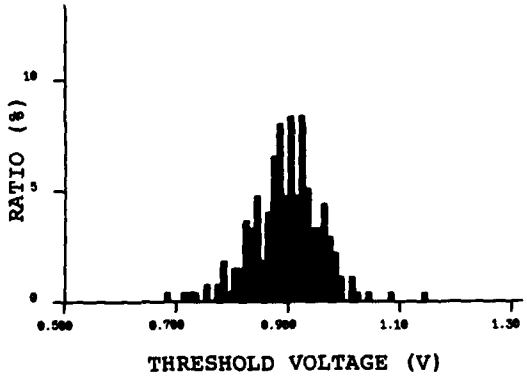


FIGURE 3

Threshold voltage distribution of N-type transistors fabricated in lamp-ZMR wafers. W= 20u.n, Leffective= 1.7um. The mean value is 0.9V and the standard deviation 61 mV.

4. DISCUSSION

We have presented above the main electrical results obtained in each of the three techniques. It must be pointed out that these results are not independant of the level of maturity of the technological device process for a given material. Since lamp-material has reached a process-compatible quality before the other two, they are a step forward on their way toward the production of SOI circuits for the microelectronics. Specific technological steps necessary. Moreover, the design constraints brought on by the relief grating have been taken into account by the circuit designers who use the defective silicon for highly doped parts of circuits. Mass-transport still has to be improved if one wants to thin the SOI film down to hundreds of angstroms. This should be useful for avoiding the "kink" effect without density loss. Another "effect" has to be reduced, that is the previously reported "shrink" of the lamp-ZMR wafers /3/. It leads to a narrowing of the defect lines. It must be quantified precisely and be reproducible in order to have it correctly corrected at the first mask level, that is the relief grating level.

On the other hand, laser- and FIPOS-SOI materials are quite new, as far as the VLSI technological process is concerned. FIPOS is very attractive since the useful material is similar to the starting bulk crystal as it does not undergo any transformation. However, the residual doping level should be reduced at least of one order of magnitude if we want to adjust the threshold voltages. Moreover, new design constraints are brought in by the fact that there is no silicon left in the access windows of the N/N+/N structure after oxidation of the porous buried layer. The interfacial quality of the oxide should also be further improved. Finally, the laser technique has to be characterized more specifically and precisely, namely concerning the residual crystalline defects (SGBs, twins, stacking faults,...) and the residual stress within the recrystallized material. Moreover, the long term stability of the laser also has to be certified. We believe, however, that this technique should find some specific applications in the field of 3-D circuits.

To summarize, we believe that it might be possible to classify the three techniques in a chronological order of applications to circuit fabrication. Lamp-ZMR material is ready for circuit development. FIPOS-material should find a medium term application, after the main residual problems indicated above will be solved. Laser-ZMR has a certified but probably long term 3-D circuit field of application and should still demonstrate its reproducibility and

crystalline quality on a large scale if it wants to be applied shortly in the VLSI 2-D technology.

5. CONCLUSIONS

We have shown that among three different techniques used in the recent years to fabricate SOI material, i.e. lamp- or laser-ZMR and FIPOS, the level of advancement is quite different in both material and technological processes. Statistical and electrical parameters obtained in lamp-ZMR batches of wafers show that this technique is ready for providing substrates for making CMOS devices. On the other hand, problems related to remaining defects in the material or to a high residual doping level still have to be solved in the other two techniques before they can provide reliable substrates.

ACKNOWLEDGMENTS

We wish to thank G. Gimine and the Pilot Line at Cnet who took an important part in the technological process. We acknowledge M. Montier and the CMS department from Thomson Semiconducteurs for his encouragement in this work.

This work was partly supported by an Esprit Project.

REFERENCES

/1/ Haond, M. and Vu,D-P., Electron. Lett. 18 (1982) 727.

/2/ Haond, M., Vu, D-P., Monroy Aguirre, A., Perret, S., Circ. and Dev. IEEE,(1987) in Print.

/3/ Haond, M., Dutartre, D. and Bensahel, D., MRS-Europe Proceedings (1985) 417.

/4/ Dutartre, D., Appl. Phys. Lett. (1986) 350.

/5/ Haond, M., Dutartre, D., Bensahel, D., MRS Symp. Proc. Vol.53 (1986) 83.

/6/ Holmstrom, R.P., Chi, J.Y., Appl. Phys. Lett., 42 (1983) 386.

/7/ Barla, K., Bomchil, G., Herino, R., Monroy, A., Gris, Y., Electron. Lett. 22 (1986) 1291.

/8/ Mermet, J.L., Achard, H., Bono, H., Joly, J.P., to be published in ESPRIT Technical Week, Bruxelles, Oct. 87.

/9/ Regolini, J.L., Dutartre, D., Bensahel, D., Karapiperis, L., Garry, G., Dieumegard, D., Electron. Lett. 23 (1987) 495.

/10/ Vu, D-P. and Pfister, J.C., Appl. Phys. Lett. 48 (1986) 50.

/11/ Vu, D.P. and Pfister, J.C., Appl. Phys. Lett. 47 (1985) 950.

/12/ Vu, D-P., Chantre, A., Mingam, H. and Vincent, G., J. Appl. Phys. 46 (1984) 1682.

/13/ Vu, D-P., Chantre, A., Ronzani, D. and Pfister, J.C., MRS Symp. Proc. Vol.53 (1986) 357.

/14/ McGreivy, D., Electron. Dev. ED-24 (1977) 730.

/15/ Davis, J.R., Glaccum, A.E., Reeson, K., Hemment, P.L.F., Electron Dev.Lett. EDL-7 (1986) 570.

/16/ Tihanyi, J. and Schlotterer, H., Solid State Electron. 18 (1975) 309.

/17/ Colinge, J.P. and Kamins,T.I., in print.

POROUS ANODISED SILICON FOR FULL DIELECTRIC ISOLATION: The Development of an n/n+/n Device Route

D Brumhead, J G Castledine and J M Keen
J M Cole, L G Earwaker, J P G Farr, P E Grzesczyk, J L'Ecuyer, M Loretto and I M Sturland*

Royal Signals and Radar Establishment, St Andrews Road, Great Malvern, Worcs WR14 3PS, UK
* University of Birmingham, P O Box 363, Birmingham B15 2TT, UK

The n/n+/n route to porous silicon has been used to produce fully dielectrically isolated silicon islands. Results are presented to show that doping of the island with residual n+ material can be avoided and that the silicon is of high crystalline perfection. The technique is shown to avoid the limitations of the original p-n technique and to be extremely promising for SOI Device applications.

1. INTRODUCTION

Progress is reported in applying oxidised porous anodised silicon for full dielectric isolation in VLSI. Early procedures [1] were based on selectively anodising p-type regions, and NTT made a 64K RAM using this technique. The method, however, had limitations with respect to island width and crystallographic perfection, and resulted in thick porous layers. Inevitably, a spike was left under the silicon islands. Other approaches involve growing good quality molecular beam epitaxial layers on porous silicon [2] or rely on a buried layer beneath the silicon wafer surface to direct the current flow laterally, under the device islands [3]. Both of these methods have disadvantages in practice.

We have explored n/n+/n enhanced lateral anodisation as an alternative route [4]. Its basis is that the suceptibility to anodising is dependent on the n-type doping level, so that higher dopant concentrations lead to lower anodising voltages and lower porous densities. Therefore n+ material can be anodised preferentially leaving a lower-doped n-type island. The advantages of this technique over using p/n selectivity are that the thickness of the porous silicon and hence the buried oxide can be decoupled from the width of the islands, there is no residual spike beneath the centre of the island, therefore the islands as formed and after oxidation are potentially defect free. A disadvantage is the need for a thin epitaxial layer with an abrupt interface to a buried n+ layer.

2. EXPERIMENTAL

A strip mask has been used to produce wafers with different structures in each of four quadrants. Island strips of width 20, 30, 40 and 60 microns have been produced with 5 micron n+ anodising entry windows between them in the four quandrants, respectively. The wafers were anodised at constant voltage in an ethanoic/HF electrolyte to convert the n+ silicon both between and under the islands into porous silicon about half a micron thick. The wafers were subsequently oxidised and analysed.

3. RESULTS

Transmission electron micrographs (TEM) of cross-sections of island-doped wafers (Figure 1) illustrate the isolation obtained, the retention of the island geometry and the good crystallographic quality of the island.

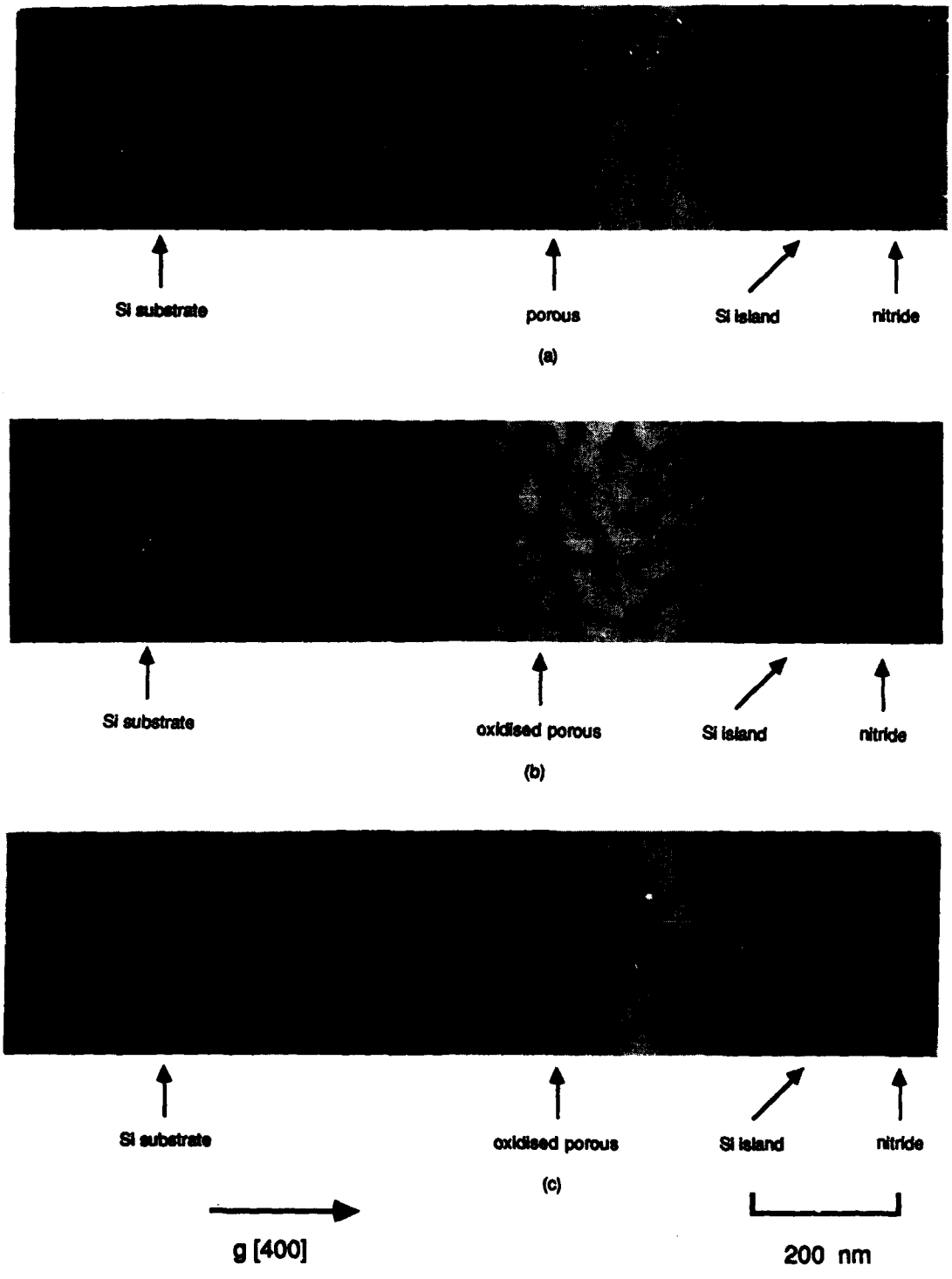


Figure 1. Cross-section TEM micrographs showing the n/n+/n microstructure following (a) anodising and (b) oxidation at 300°C for 1 hour and 800°C for 2 hours, (c) 300°C for 1 hour, 800°C for 1 hour and 1090°C for 4 minutes followed by an anneal in dry nitrogen for 1 hour. Note the absence of defects in the Si island and the sharpness of the interface between the Si island and the porous Si.

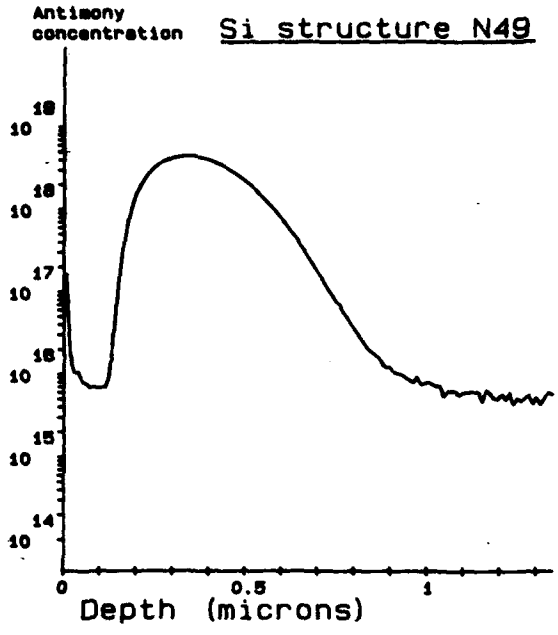


Figure 2(a). SIMS profile showing the initial dopant distribution through a vertical section of the island structure.

The non-uniform structure of the as-anodised porous silicon is related to the implant profile which is well-revealed by SIMS analyses (Figure 2a). The pore size and structure is determined by a complex relationship between dopant concentration, electrolyte composition and the anodising current density. By careful choice of conditions, the doping levels and current density effects can be used to counteract each other in order to achieve a much better uniformity than that shown.

The steepness of the profile determines the roughness of the interfaces between the silicon island/porous silicon, and substrate/porous silicon. On oxidation, the porous silicon, together with some of the silicon at the back of the island and at the substrate interface, is converted to SiO_2. This is shown in Figure 1 (TEM) and Figure 3 (Rutherford Backscattering).

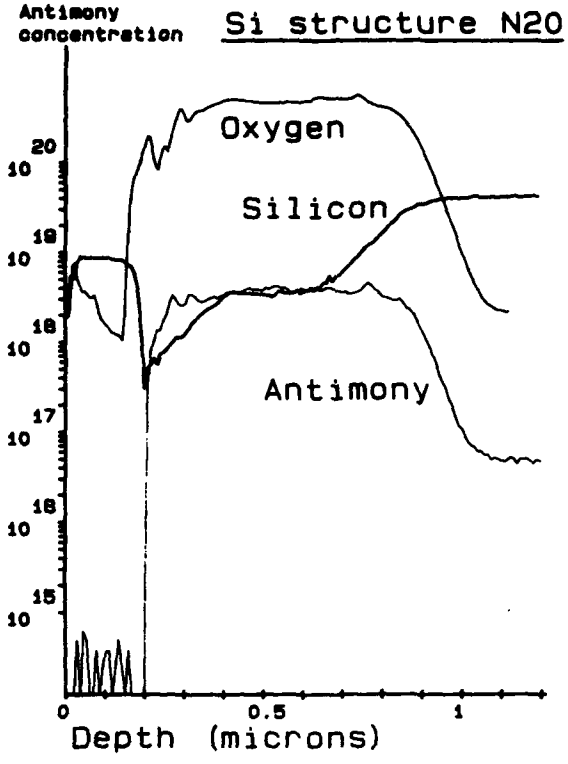


Figure 2(b). SIMS profile showing dopant distribution after the oxidation of the porous silicon.

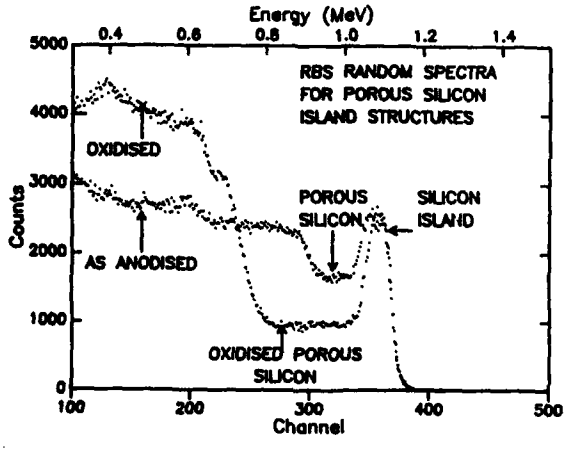


Figure 3. Rutherford backscattering (RBS) spectra from randomly oriented samples of as-anodised and anodised-plus-oxidised (1 hour at 300°C plus 2 hours at 800°C) porous silicon. The spectrum from the oxidised porous sample indicates fully oxidised SiO_2 with some of the back face of the island also being oxidised.

If the anodising has been done correctly, the oxidation of the whole of the back of the island is uniform. Due to the steepness of the impurity profile and the low oxidation temperatures involved, an oxide barrier can be created which makes it possible to avoid doping the back of the island (Figure 2b).

4. CONCLUSIONS

Very lightly doped islands have been produced with sharp interfaces, good island morphology and low defect densities. The n/n+ route has been shown to overcome the limitations associated with the original p/n porous silicon technique. It is therefore, an extremely promising approach to SOI device production.

ACKNOWLEDGEMENTS

This work was partially funded by the UK Alvey programme.

The assistance of Plessey Research (Caswell) Ltd and the GEC Hirst Research Centre in the preparation of these samples and of Mr Blackmore (RSRE) for the SIMS analyses is gratefully acknowledged.

REFERENCES

[1] K Imai, Solid State Electronics, Vol 24 (1981) pp 159-164.

[2] G Konaka, M K Tabe and T Sakai, Appl Phys Lett 41 (1983) 86.

[3] J Benjamin, J M Keen, A G Cullis, B Innes and N G Chew Appl Phys Lett 49 (12) 22 September 1986, 716.

[4] R P Holmstrom and J Y Chi, Appl Phys Lett 42 1983 386.

© Copyright at HMSO, London, 1987

STACMOS: A BASIC 3-DIMENSIONAL CMOS PROCESS

R. Buchner, K. Haberger, P. Seegebrecht and P. Panish

Fraunhofer Institut fur Festkörpertechnologie
Paul-Gerhardt-Allee 42
D-8000 Munich 60, West Germany

MOS Transistors have been fabricated in two independent active device layers, the second of which has been formed through laser recrystallization of a thin polysilicon layer. The effect of the fabrication process on the devices in the silicon substrate has been investigated and characterized through electrical measurements.

1. INTRODUCTION

Integrated circuit semiconductor development is characterized by constantly increasing device packing density. Until recently this has been achieved by reducing the lateral device dimensions while increasing the chip area. The magnitude of reduction in device dimensions has proved to introduce significant technological problems. Another possibility for increasing the level of integration is the utilization of one or more additional device layers. In addition to reducing the interconnection length such a tactic allows the realization of completely new circuit concepts as well as making possible the utilization of mixed technologies.

In order to investigate the effect of recrystallization of the upper layer of a 3-dimensional circuit on the devices in the underlying layer a 3D-CMOS process has been developed.

2. PROCESS

Since the primary purpose of this investigation is to determine the effects of the recrystallization process, the technology has been conceptualized to be as simple as possible. Two active device layers have been fabricated, the first of which is in the monocrystalline silicon substrate, and the second which is fabricated in a thin recrystallized polysilicon layer. This arrangement suited itself to a CMOS circuit structure in which the NMOS and PMOS devices are fabricated in their own layers respectively thus simplifying the process, though it is in principle not necessary. Due to the fact that arsenic is the most thermally stable element in silicon it was decided that the n-channel devices would be fabricated using As doped source-drain regions in the monocrystalline silicon substrate.

The 3D process developed requires 10 mask steps and utilizes virtually exclusively standard semiconductor processes. Starting material for the process is 3" p-type (100) silicon wafers. The first phase of the process is the utilization of a standard poly-gate MOS process to fabricate the n-channel devices using a 500 A thick gate oxide with a minimum feature size of 4 um. The source-drain regions are As doped using ion implantation. At the completion of this phase the surface exhibits significant topography. It is known that such nonplanar features have undesirable effects on the recrystallization process resulting in nucleation sites for grain boundary formation [1]. These effects are minimized by applying a planarization process which serves simultaneously to insulate the first device layer. A two stage planarization process is used. In the first stage a LPCVD oxide is deposited and then patterned to fill the depressions in the

source-drain areas. A low temperature oxide (LTO) is used for which the etch rate in buffered HF is roughly three times that of a thermal oxide thus allowing processing without special etch stops. With the proper mask design this process allows a nearly planar surface to be obtained. The remaining irregularities are minimized in a reflow planarization step in which a 0.1 um oxide layer is deposited followed by 0.4 um of Phosphorous-Silicate-Glass (PSG). After reflow a 0.2 um oxide layer is deposited to serve as a diffusion barrier.

After completion of the devices in the first layer a 0.5 um thick layer of polysilicon is deposited, doped, and covered with an 850 A thick LPCVD oxide capping layer. The polysilicon is recrystallized over the entire surface with the help of an argon laser. This is necessary in order to avoid excessive absorbtion in the underlying monocrystalline substrate. The laser power, beamwidth, scan speed and substrate temperature are 12 W, 70 um, 3 cm/sec and 500 °C respectively. Modification of the melt-zone temperature profile is achieved through beam forming as shown in Figure 1. Following recrystallization the silicon layer is divided into individual islands to allow for contacts to the devices in the substrate layer.

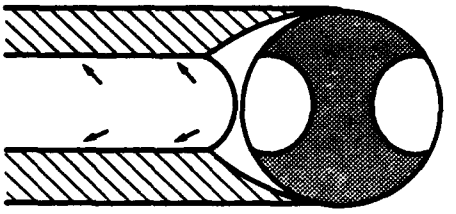


Fig. 1: Beam form of the argon laser.

P-channel polysilicon gate MOS transistors with a 500 A thick gate oxide are formed in the recrystallized silicon layer using a modified standard MOS process. A substrate contact for the devices in the second layer is included since it has been found that the characteristics of devices with a floating substrate are different than those at a fixed potential. Additionally the fixed substrate allows for minimization of coupling effects between the two layers.

The contact window opening proceeds using a combined dry/wet etch process which takes advantage of the high selectivity of wet chemical etching while obtaining low levels of lateral underetching.

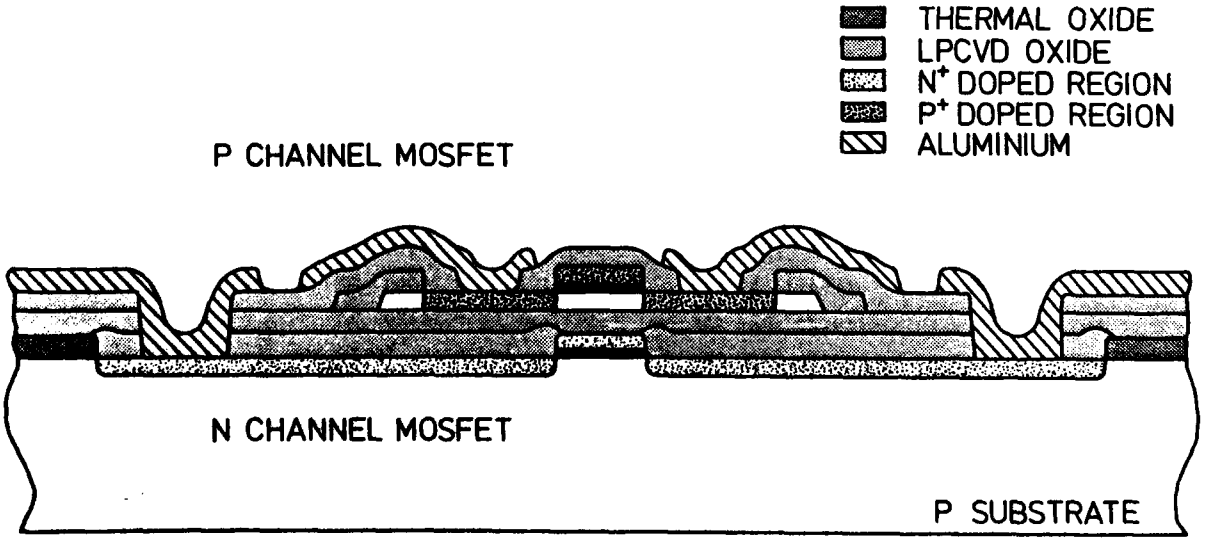


Fig. 2: Schematic cross section of a 3D MOS device.

3. RESULTS

A schematic cross-section of a 3D device consisting of two MOS transistors is shown in Figure 2. A SEM photograph of the corresponding fabricated structure is shown in Figure 3.

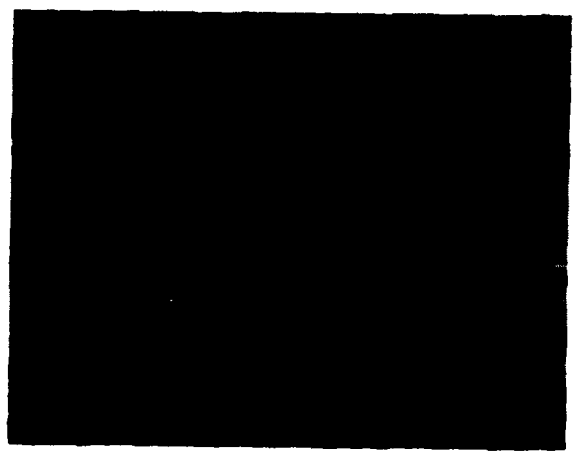


Fig. 3: SEM photo of a 3D device.

As stated earlier the main goal of this investigation was a determination of the effect of the recrystallization procedure on the underlying devices in the monocrystalline substrate. Characteristic curves typical of the MOS devices in both layers are shown in Figures 4 and 5. Although the threshold voltage of the n-channel devices is slightly negative this may be easily compensated with an appropriate channel implantation. The p-channel transistors exhibit a relatively high threshold voltage of -12 V and a mobility of roughly 140 cm^2/Vs. This may be attributed to high levels of surface roughness on the boundary layer at the silicon/gate-oxide interface as has been seen with SEM investigation. The source of this roughness may lie in the low scan speed used during recrystallization.

The measurements show that the recrystallization process has a limited effect on the parameters of the devices in the substrate layer. Although the process has no significant effect on the mean values of the threshold voltage and mobility, the parameter scattering increases by roughly 50%. After completion of the second active layer the mean parameter values have shifted as well. It is possible that this shift is caused by the resultant mechanical stress between the

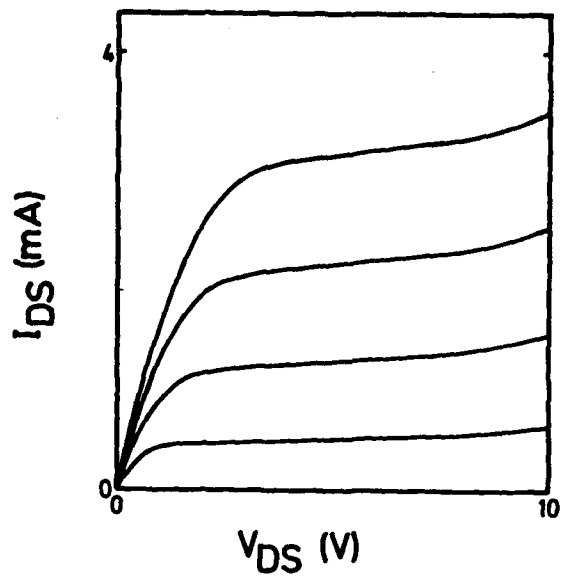


Fig. 4: Typical characteristic curve of a n-channel MOS transistor.

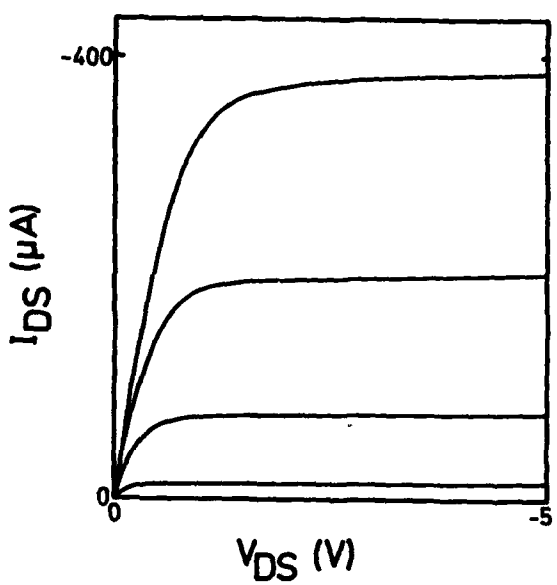
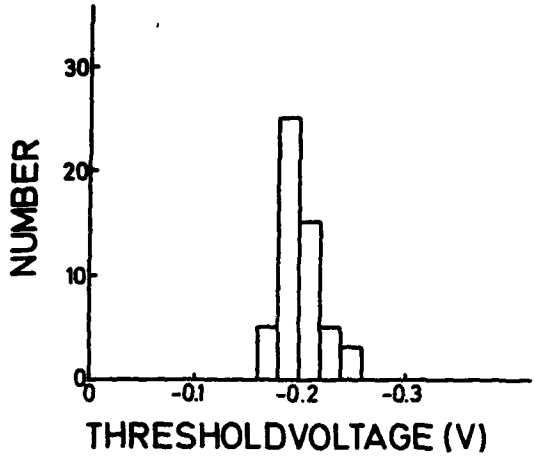
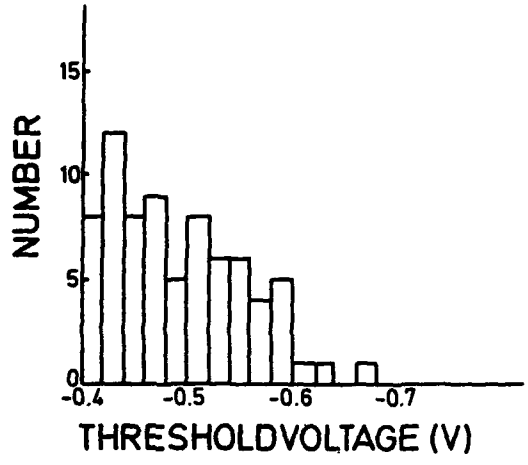
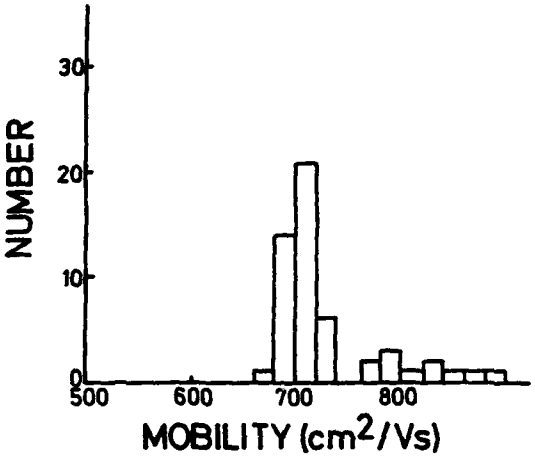


Fig. 5: Typical characteristic curve of a p-channel MOS transistor.







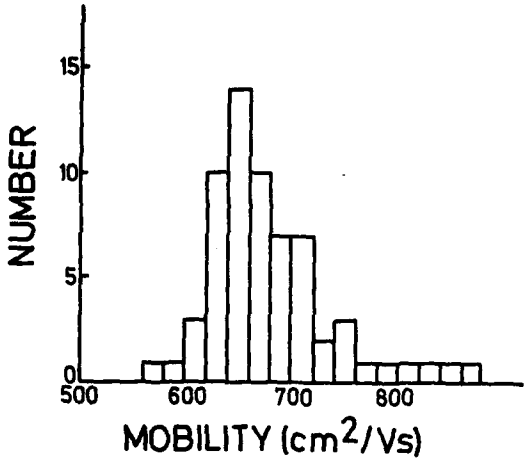


Fig. 6: Measurement distribution of the n-channel transistors after completion of the first layer.

Fig. 7: Measurement distribution of the n-channel transistors after completion of the second layer.

individual layers. The histograms showing these effects are displayed in Figs. 6 and 7.

4. SUMMARY

This investigation has shown that the fabrication of a second active device layer has only a minor effect on the characteristics of the devices in the underlying layer. The process presented here has served as the basis for the development of a 2 um design rule 3D-CMOS process.

ACKNOWLEDGEMENTS

The authors are grateful for the technical support of M. Forster, A. Heidenreich and B. Brandt. This work was supported by the West German Ministry of Research and Development.

REFERENCES

[1] Miyao, M., Ohkara, M., Takemoto, I., Tamara, M., and Tokuyama, T., Appl. Phys. Lett., 41(1), July 1982

VOLUME INVERSION IN SOI MOSFETs WITH DOUBLE GATE CONTROL: A NEW TRANSISTOR OPERATION WITH GREATLY ENHANCED PERFORMANCE

F. BALESTRA, S. CRISTOLOVEANU, M. BENACHIR, J. BRINI and T. ELEWA

Laboratoire de Physique des Composants à Semiconducteurs (UA-CNRS),
ENSERG/INPG, 23 Av. des Martyrs, 38031 Grenoble, FRANCE

Silicon-On-Insulator transistors are used with a double gate control. By this way, a fully inverted silicon film (interface and film volume) is obtained. This method allows us to greatly enhance the device performance, in particular the subthreshold swing, transconductance and drain current. Simulated and experimental characteristics on SIMOX structures are analysed to study the new device.

1. INTRODUCTION

Silicon-On-Insulator materials present many advantages compared with the bulk silicon VLSI technology: lateral isolation, lower parasitic capacitance and power, higher speed, reduced short channel effects, radiation tolerance, ...

In this communication, the special multi-interface configuration of SOI structures is used to obtain a new device based on volume inversion. The theoretical analysis is achieved with a "home-made" computer program (ISIS) which gives the solution of the Poisson equation in multilayer structures [1], and the experiment is carried out on SIMOX devices.

2. SIMULATION

The physical principle of the device is shown in Fig.1. Surface inversion channels can be activated either at the top interface or at the back interface using the normal gate V_{G1} or the secondary gate V_{G2} (bulk Si substrate) respectively. We choose to simultaneously bias both gates ($V_{G2} = K V_{G1}$), where the coefficient K accounts for the differences in thickness and threshold voltage between gate oxide and buried oxide ($K \approx 10$).

If the film is thick or highly doped, there is no overlap of the two depletion regions and the inversion channels grow almost independently. For example in Fig. 1a, the film is slightly depleted for $V_{G1} = 2V$ and only a low coupling appears between the two conducting channels.

A different behaviour (Fig.1b), caused by the coupling of the two interfaces, occurs in films with normal thickness (< 0.2 μm) and low doping (a few 10^{15} cm^{-3}). For $V_{G1} = -1$ V, the whole silicon film is in accumulation. For higher V_{G1}, the potential increases at the interfaces and in the film volume, from depletion to weak and strong inversion. For $V_{G1} \geq 0.6$ V, the potential shift exceeds $2\Phi_F$ in every regions and all the film is in strong inversion. We propose to call this new device the "volume-inversion MOSFET" (VI-MOSFET).

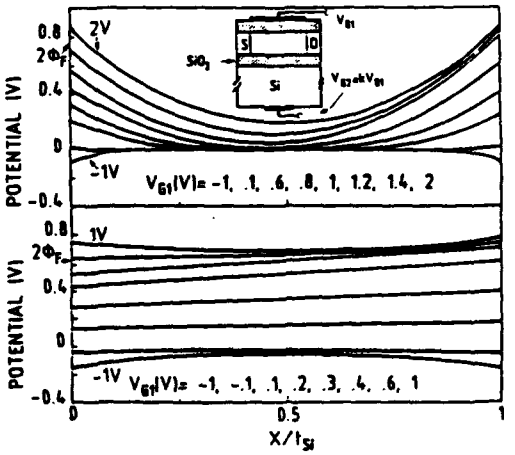


Figure 1

Potential profiles inside the silicon film for uncoupled (a : doping $N_a = 4.10^{16}$ cm^{-3}, film thickness $t_{Si} = 300$ nm) and coupled (b : $N_a = 3.10^{15}$ cm^{-3}, $t_{Si} = 100$ nm) interfaces. (27 nm and 380 nm thick oxides).

The behaviour of the VI-MOSFET is governed by minority carriers, which now are no longer confined at an interface. There are significant advantages: greatly increased number of minority carriers, reduced influence of surface scattering and interface defects, use of the volume which is much thicker than a surface inversion layer and has higher carrier mobility. These special features lead to a great improvement in current value, subthreshold slope, transconductance and speed.

In Fig.2 are compared the current-voltage characteristics of an N-channel VI-MOSFET (K = 10) with those of a normally operated MOSFET (K = 0, i.e. inversion layer at the top surface only). The current and transconductance variation versus V_G are calculated using a mobility profile suggested by transport measurements [2] : 1200 cm^2/Vs in the center of the film, 500 cm^2/Vs at the front interface and 400 cm^2/Vs at the back interface.

The subthreshold swing (Fig.2a) of the VI-MOSFET (29 mV/decade, curve 1) is excellent, compared with 66 mV/decade for the normally operated MOSFET (curve 2), and clearly goes far below the theoretical limit of the normal MOSFET (≈60 mV/decade).

In strong inversion (Fig.2a) the current of the VI-MOSFET (curve 1) exceeds by a factor 3 at V_{G1} = 1.2 V that of the normal MOSFET (curve 2). This is due to (i) the increase in the total number of carriers, (ii) the improvement of the subthreshold swing and (iii) the transconductance overshoot of the VI-MOSFET (Fig.2b, curve 1). Indeed, the transconductance is clearly enhanced for K = 10 (curve 1) in comparison with K = 0 (curve 2). The maximum field effect mobility (1050 cm^2/Vs), corresponding to the transconductance maximum, is close to the carrier mobility in the film volume.

The comparison with a linear variation of carrier mobility (without peak mobility in the film center) between the two interfaces (curve 3 : K = 10 and curve 4 : K = 0), emphasizes the importance of the volume mobility. It is interesting to note that even for K = 0 (curve. 2 and 4) the mobility profile is important, since a volume inversion still exists in a narrow region close to the front surface.

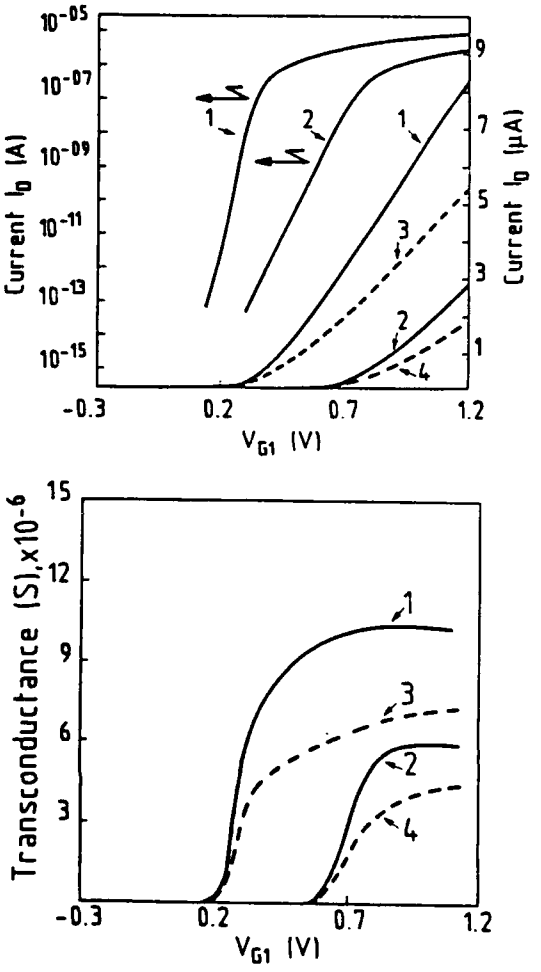


Figure 2

Simulated I-V characteristics (a) and transconductance (b) for K = 10 (curve 1) and K = 0 (curve 2), for the device of Fig.1b and a bell shaped mobility profile. The curves 3 (K = 10) and 4 (K = 0) are obtained with a linear mobility profile. For the sake of simplicity, the mobilities are supposed to not depend on V_G.

Fig. 3 shows the current-voltage characteristics of N-type depletion mode transistor. The drain current and the transconductance are compared for the normally operated MOSFET (K = 0) and for the MOSFET with volume-accumulation (VA-MOSFET) with K = 10. Volume accumulation is more easy to obtain than volume inversion, because it is a "natural" behaviour of the silicon film for a normal transistor in flat band situation. Therefore, the gains are slightly lower for the VA-MOSFET than for the VI-MOSFET. Nevertheless, the use

of a double gate control greatly improves the important device parameters. Indeed, we can observe in Fig.3 a decrease of the subthreshold swing (67 mV/dec for K=0 (curve 2) and 34 mV/dec for K=10 (curve 1)), an increase of the transconductance (90 %) and current value (75 %) for the VA-MOSFET.

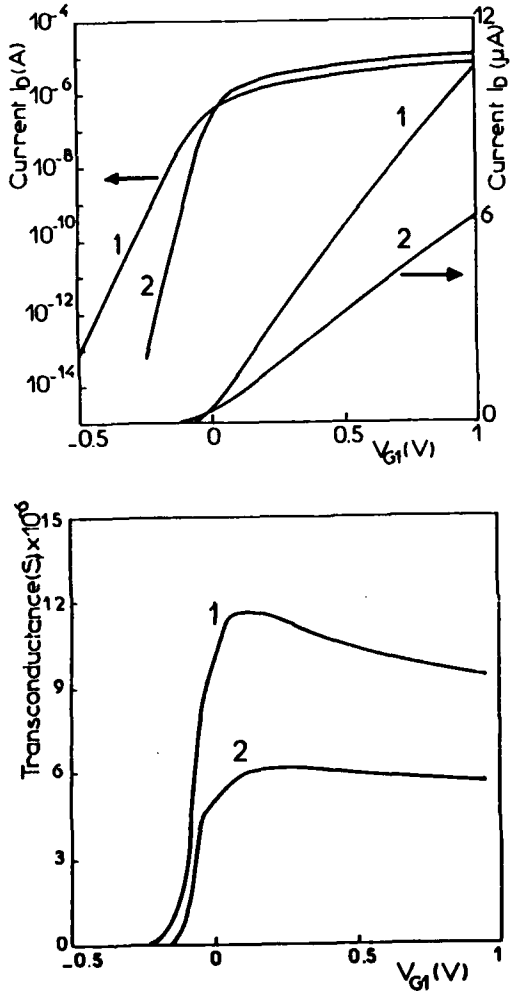


Figure 3

Simulated I-V characteristics (a) and transconductance (b) for K = 10 (curve 1) and K = 0 (curve 2) for a depletion mode transistor with the same technological parameters than in Fig.1b (doping : $N_d = 3.10^{15}$ cm^{-3}).

3. EXPERIMENTS

All these optimistic expectations are indeed verified. Experimental evidence has been obtained with P-channel MOSFETs fabricated with standard technology on a low-doped SIMOX substrate (dose $1.8x10^{18}$ O^+/cm^2, energy 200 keV, annealing above 1300°C). The thicknesses are about 200nm for the film, 27 nm for the gate oxide and 380 nm for the burried oxide. The characteristics of 0.8 μm long transistors are given. Fig.4 a and b clearly show the great gains in subthreshold swing (70 mV/decade for K=0 (curve 1) and 29.5 mV/decade for K=10 (curve 2)), transconductance (80%) and current value which increases by a factor 2.8 at V_{G1} = -2 V, by using a double gate control. The leakage currents of the VI-MOSFET is very low because carrier accumulation occurs simultaneously at both interfaces. Similar improvements have been obtained on N-channel devices.

The enhancements for the case of depletion mode transistors with volume accumulation are the same than those of the simulated characteristics.

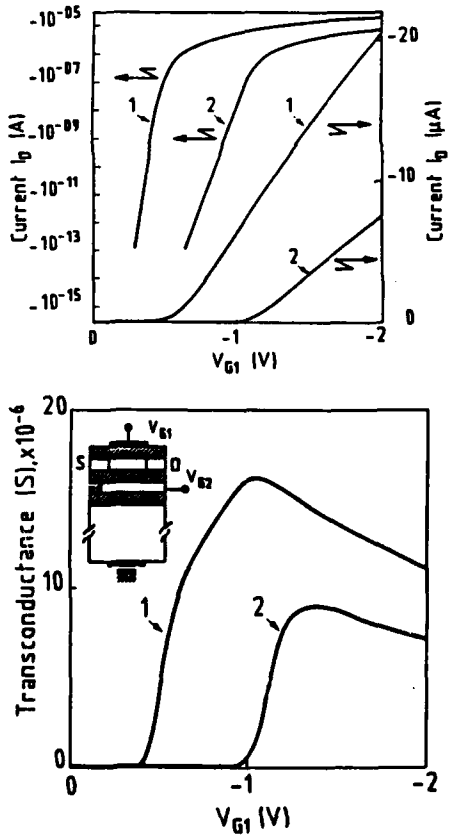


Figure 4

Experimental I-V characteristics (a) and transconductance (b) for a P-channel transistor made on SIMOX.

For practical applications of VI-MOSFETs (or VA-MOSFETs), to overcome the limitation due to the use of the substrate as a unique secondary gate, the double SIMOX structure obtained by two oxygen implants [3] can be proposed (see insert of Fig.4b). By this way, a silicon layer, lying between the two burried oxides, is formed and acts indeed as a gate. Segregation of the secondary gates of various VI-MOSFETs is made by oxidizing the unuseful portion of this film (by simple adjustement of the oxygen implantation) or by etching.

On the other hand, to decrease the voltages applied on the back gate, the difference in thickness between the gate and burried oxide can be reduced. An other solution is to scale down the biases of VLSI circuits.

4. CONCLUSION

In conclusion, the transistor performances have been improved using the new principle of double gate control of VI-MOSFETs or VA-MOSFETs. The experimental enhancements (current, transconductance and subthreshold swing gains) are in agreement with our theoretical calculations.

ACKNOWLEDGMENTS

The authors are indebted to Dr. A.J. Auberton-hervé (LETI, Grenoble) and to Dr. J. Davis (British Telecom, Ipswich) for providing SIMOX devices.

REFERENCES

[1] F. Balestra, J. Brini and P. Gentil, Solid State-Electron. 28, 1031 (1985).

[2] S. Cristoloveanu, S. Gardner, C. Jaussaud, J. Margail, A.J. Auberton-Hervé and M. Bruel, submitted to Journ. Appl. Phys. (1987).

[3] G.K. Celler, J.L. Batstone, K.W. West, P.L.F. Hemment and K.J. Reeson, IEEE SOS/SOI Workshop (Captiva, USA), Oct. 1986.

A NEW TYPE OF HIGH PERFORMANCE DEVICE FOR VLSI DIGITAL SYSTEM

XU XIAO-LI, TONG QIN-YI, XONG HE-MING

Microelectronics Center, Nanjing Institute of Technology
Nanjing 210018, China

This paper presents a high performance Complementary Buried Channel FET device isolated by high quality silicon dioxide layer using Silicon wafer Direct Bonding technology (SDB/CBCFET). The structure and operational principle of this device is discussed. The properties of an improved SDB process is investigated. By means of 2D numerical simulation, effects of interface charge density of bonding interface and SOI layer-SiO2 interface on threshold voltage and the threshold voltage shift in submicron geometry are analysed. The performance of submicron SDB/CBCFET device and circuits is evaluated. The results indicate that SDB/CBCFET device is superior to bulk CMOS and SOI/CMOS in speed, switching energy, complexity, reliability and small size effects as device size decreases into submicron dimension.

1. INTRODUCTION

The dominant technology used in modern VLSI digital system is bulk CMOS. The development of bulk CMOS has been supported primarily by rapidly decreasing feature size in the circuits. However, as bulk CMOS technology advances into submicron dimension, several limitations have become apparent: complex process sequences, significant short channel effects and serious interaction between neiboring devices. These negative effects have limited CMOS to improve its performance by further scalling. Silicon-On-Insulator CMOS technique offers an attractive alternative by providing simple device fabrication sequences, improved short channel effects and no latch up problems. However, the poor quality of SOI substrate by conventional techniques such as SIMOX, laser anealling etc. and SOI versions of corresponding bulk CMOS device have limited better use of SOI potential advantages. Moreover, the negative effects of CMOS device can not be effectively suppressed. In recent report (1), we investigated a high quality SOI substrate technology (SDB) which shows that the quality of SOI layer and the underlying SiO2 layer is not degraded from its original "bulk" quality. Based on the characteristics of this technology, we propose a new complementary buried channel device which has higher bulk mobility, much smaller short channel effects and higher performance than bulk CMOS and SOI/CMOS.

In this paper, the structure and operational principle of SDB/CBCFET is described, its small size effects and its device/circuits performance are discussed by means of 2D numerical simulation and analysis models.

2. DEVICE STRUCTURE

SDB/CBCFET is characterized as a compound structure of MESFET and MOSFET (see Fig.1). In its lateral direction, SDB/CBCFET has the same doping type for source, gate and drain which is similar to MESFET, while in its vertical direction, SDB/CBCFET employes poly-Si and SiO2 layer as its gate which is identical to that of MOSFET. By adjusting doping concentration of both poly-Si

gate and channel region, as well as selecting proper SiO2 thickness, the p channel and the n channel device can both be held in deep depletion at zero gate bias and normally off characteristics both for n channel and p channel device can be obtained. Therefore, an n channel device and a p channel device of this type structure can form a basic complementary inverter.(see Fig.

2.1. Properties of SDB Substrate

An SDB technology has been developed. The bonding process adopted in our work involves the following main process sequences: the two mirror polished wafers oxided in wet oxyen, the wafers treated in acid solution, the wafers contacted face to face put into a N2 ambient at 1050C° for about 1 hour forming the SOI substrate. The microstructure and electrical properties of the SOI substrate using this SDB technology have been extensively studied by TEM and DLTS analysis. The experiment results show that dislocations are concentrated on the back side of the substrate and no additional defects have been developed within 80um from Si02-Si02 bonding area. Therefore the interface charge density between SOI layer and underlying Si02 is no more than the interface charge density between conventional bulk silicon and thermal oxidation layer. Both hole and electron mobilities are measured using Van Der Paw method. The results show that hole and electron mobility of the SOI substrate are 369cm /vs and 1079cm /vs respectively which is almost identical to that in original bulk silicon. Since the quality of SOI substrate by SDB technology is not degraded from its original "bulk" silicon quality, this SOI substrate is very suitable for CBCFET device structure.

2.2. Features of CBCFET

In contrast to conventional CMOS, SDB/

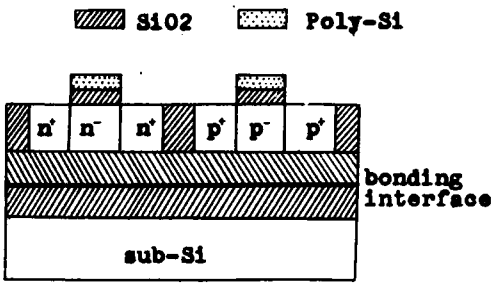


Fig.1 : Cross-Sectional view of SDB/CBC FET configuration

CBCFET has no p-n junctions at source and drain region. Thus, a significant reduction of small size effects is achieved resulting in much less threshold voltage shift, improved punchthrough resistance, and higher performance.

Channel carrier mobility in SDB/CBCFET is high due to its buried channel nature. Furthermore, the low channel doping concentration and the low threshold voltage is responsible for high speed and low power consumption of SDB/CBCFET due to its low logical voltage swing, low supply voltage and high bulk mobility.

3. DEVICE CHARACTERISTICS

The 2D numerical simulation results show that the interface characteristics of SOI layer-underlying Si02 interface deeply affect device properties. If the interface charge density is large and the thickness of SOI layer is thin, for the n channel device, normally-off operation changes into normally-on operation and for the p channel device, it is in deep depletion status which increases its threshold voltage greatly. Fig.2A shows threshold voltage vs. SOI layer-underlying Si02 interface charge density for n channel device with various SOI layer thickness. Fig.2B shows the relation between n channel threshold voltage

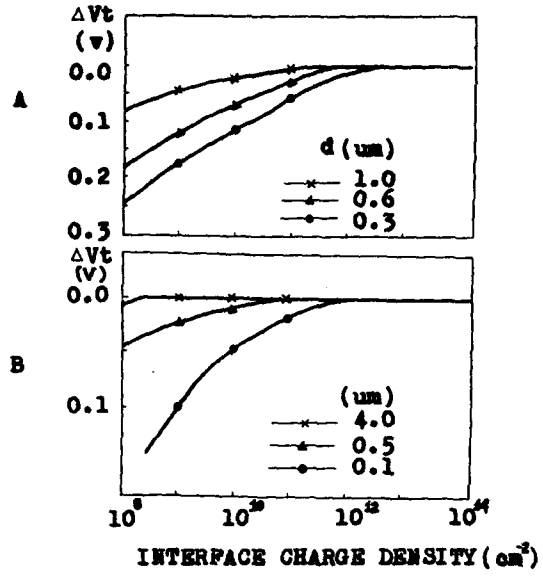


Fig.2
N channel threshold voltage shift vs. interface charge density with 0.5um channel length, 0.78V flat band voltage and 1V supply voltage for A: SOI layer-underlying SiO2 interface with various SOI thickness, B: SiO2-SiO2 bonding interface with various SiO2 thickness.

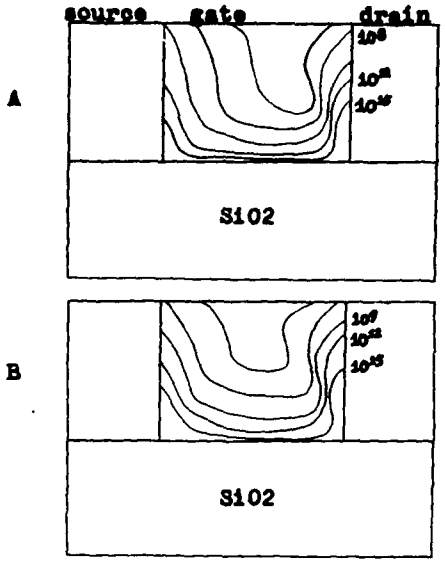


Fig.3 : 2D doping concentration distrabution of n channel device (at o.3V threshold voltage) (A) and p channel device (at -0.3V threshold voltage) (B) with 0.5 um channel length, 0.78V flat band voltage, 5x10 interface charge density and 1.0V supply voltage.

and bonding interface charge density for various underlying SiO2 layer thickness. From Fig.2B we can conclude that as long as the thickness of underlying SiO2 is larger than 0.5um, the bonding interface charge density will almost not affect the threshold voltage of SDB/CBCFET.

Fig.3A&B shows 2D doping concentration distrabution of n channel device (at 0.3 V threshold voltage) and p channel device (at -0.3V threshold voltage). It is found that both devices show normally-off operation and small punchthrough effects.

Threshold voltage shift as a function of effective channel length for n channel device is shown in Fig.4. The amount of threshold voltage shift of SDB/CBCFET determined by 2D numerical simulation is compared with that of bulk CMOS and that of SOI/CMOS (2). It is shown that for the submicron device, the threshold voltage

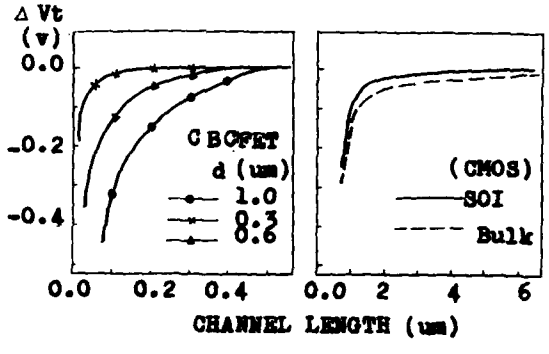


Fig.4 : Threshold voltage shift of n channel device vs. channel length, with 0.78V flat band voltage, 5x10 interface charge density and 1.0V supply voltage.

shift of n channel device of SDB/CBCFET is much smaller than that of CMOS device structure. The 2D numerical analysis explains the physical origin of the difference between bulk CMOS, SOI/CMOS and SDB/CBCFET. For SDB/CBCFET, absence of

TABLE 1

	L (um)	d (um)	Nd (cm^{-3})	un/up (cm^2/vs)	Vtn/Vtp (v)	Vd (v)	T (ps)	E (fJ)
Bulk CMOS	0.65	5	$1x10^{15}$	453/212	0.7/-0.5	5	460	> 100
SOI/CMOS	0.8	0.5	$1x10^{16}$	680/280	0.4/-0.4	5	93	> 50
SDB/CBCFET	0.5	0.3	$1x10^{16}$	1079/369	0.3/-0.3	1	5.5	3.96

Where L: channel length; d: thickness of Si layer; Nd: doping concentration of active region; un&up: mobility of n channel & p channel device; Vtn&Vtp: threshold voltage of n channel & p channel device; Vd: supply voltage; T: delay time per gate of ring osillator; E: switching energy.

p-n junction at source and drain, small channel depletion charge due to low channel doping and thin SOI layer contribute to much less threshold voltage shift.

4. CIRCUIT PERFORMANCE

In order to show the important characteristics of SDB/CBCFET device and circuits, we have designed a ring osillator composed of SDB/CBCFET device, optimised design parameters and evaluated ring osillator performance (see Table 1). The results indicate that the SDB/CBCFET is superior to bulk CMOS (3) and SOI/CMOS (4) in speed, switching energy and power dissipation as device size advances into submicron dimension.

5. CONCLUSION

A new high performance complementary buried channel device isolated by high quality SiO2 layer using SDB technology has been presented. The characteristics of the device and performance of the circuits have been evaluated. Following results are obtained: 1) simple process sequences, high carrier mobility and small interface charge density of SOI layer by SDB technology, 2) small threshold voltage shift for both n channel and p channel device, 3) high speed, low power, high reliability and complexity of this complementary device. The results indicate that SDB/CBCFET is a potential candidate for very high performance VLSI digital system.

REFERENCES

(1) Li Hui, Tong Qin-Yi, etal., INFOS 87, C.4.4.

(2) Michael P.Brassington, et al., IEEE Trans. Electron Device, vol. ED-32, p1858,1979.

(3) Theodore I. Kamins et al., IEEE Trans. Electron Device Letters, vol. EDL-6, p617, 1985.

(4) A.J.Auberton, et al., IEDM 84, 34.5.

Session C2.2

MOS Modelling
I

Chairman: G. Baccarani

Tuesday, September 15, 1987

COMPARISON OF LONG- AND SHORT-CHANNEL MOSFET'S CARRIED OUT BY 3D-MINIMOS

M. Thurner, S. Selberherr

Institut für Allgemeine Elektrotechnik und Elektronik
Technische Universität Wien
Gusshausstraße 27-29, A-1040 Wien, AUSTRIA

An accurate three-dimensional simulation program for MOSFET devices has been developed by extending MINIMOS (vers. 4) in 3D. The physical model is based on the 'hot-electron-transport model', which includes the Poisson equation, the continuity equations and a selfconsistent set of equations for the currents, mobilities and carrier temperatures. The standard finite difference discretization and the SOR (successive over relaxation) method are utilized to reduce computational time and memory requirements. Adaptive grid refinement is used to equidistribute the discretization errors. Three-dimensional effects like threshold shift for small channel devices, channel narrowing and the accumulation of carriers at the channel edge have been successfully modeled. Our comparison of several MOSFET's make clear that three-dimensional calculations are most important for accurate device modeling.

1 Introduction

The shrinking dimensions of the elements of IC's require for accurate simulation suitable device models in physics and mathematics. The two-dimensional device simulations performed in earlier times described the electrical characteristics for large transistors well but the advanced VLSI technology led to serious problems in modeling such devices and therefore a great demand appeared for 3D simulations.

The three-dimensional effects in MOSFETs like the increasing threshold voltage and the shift of the breakdown voltage caused by the finite channel width are not taken into account by the two-dimensional simulations [1]; the 2D programs are meanwhile state of the art. Accurate investigations of the previously stated effects and the knowledge of increased current densities under certain bias conditions at the channel edge are important not only for studying the electrical device characteristics but also for aging effects [2]-[3]. Therefore we have extended the two-dimensional MINIMOS to a three-dimensional simulation program. A realistic physical model and suitable mathematical algorithms have been developed to simulate the previously stated three-dimensional effects.

We shall report in Chapter 2 about the physics and the mathematics on which the simulations are based.

The results of our simulations carried out by 3D MINIMOS are reported in Chapter 3 and will be discussed there, too. We shall show that the three-dimensional simulations are indispensable for the advance from VLSI- to ULSI technology.

2 The Physical Model and the Mathematical Algorithms for the Three-Dimensional Simulation

The physical model for the simulation program is given by the Poisson and the continuity equations and the drift-diffusions model for the carrier current densities.

$$div\ grad\ \psi = \frac{q}{\varepsilon}(n - p - C) \tag{1}$$

$$div\ J_n = qR \tag{2}$$

$$div\ J_p = -qR \tag{3}$$

$$J_n = -q\mu_n(n\ grad\ \psi - grad\ (U_{t_n} n)) \tag{4}$$

$$J_p = -q\mu_p(p\ grad\ \psi + grad\ (U_{t_p} p)) \tag{5}$$

The Poisson equation (1) will always be solved fully

three–dimensionally; the continuity equations (2) and (3) at the first level of sophistication are solved two–dimensionally in the middle of the channel width. The carrier distribution in the whole volume will be calculated by the assumption of negligible current flow in the third direction $J_{n_z} = J_{p_z} = 0$. Assuming the validity of Boltzmann statistics the previous statement is equal to constant quasi Fermi levels in the direction of the channel width

$$\frac{\partial \varphi_n}{\partial z} = \frac{\partial \varphi_p}{\partial z} = 0$$

So we can write

$$n_{x,y,z} = n_{x,y,\frac{w}{2}} \cdot exp(-\frac{1}{U_t} \cdot (\psi_{x,y,\frac{w}{2}} - \psi_{x,y,z})) \quad (6)$$

$$p_{x,y,z} = p_{x,y,\frac{w}{2}} \cdot exp(+\frac{1}{U_t} \cdot (\psi_{x,y,\frac{w}{2}} - \psi_{x,y,z})) \quad (7)$$

The index $\frac{w}{2}$ denotes the middle of the channel width.

The second level of sophistication is obtained by assuming negligible current flow in the third dimension for the majorities and solving the continuity equation for the minorities fully three–dimensionally.

The third level is the fully three–dimensional solution of the continuity equations for both the minorities and the majorities.

For solving the previously specified set of equations we apply for discretization the standard finite difference method. The grid generation will be performed by an automatic mesh refinement algorithm which equidistributes the discretization error.

The linearized equations are essentially solved with an iterative algorithm. In our case we apply the SOR (Successive Over Relaxation) method. The general iterative algorithm:

$$B \cdot x^{(n+1)} = (B - A) \cdot x^{(n)} + b$$

is solved with the matrix $B = (\frac{1}{\omega} \cdot D - L)$. D is the diagonal part of A which is transformed to the unity matrix while L is the lower triangular part of A. With respect to the special linearization method one unknown reduces to:

$$\begin{aligned} x_i^{(n+1)} = (1-\omega) \cdot x_i^{(n)} + \omega(b_i - \\ - x_{i-1}^{(n+1)} \cdot a_{i-1} - x_{i-NX}^{(n+1)} \cdot a_{i-NX} - \\ - x_{i-NXY}^{(n+1)} \cdot a_{i-NXY} - \\ - x_{i+1}^{(n)} \cdot a_{i+1} - x_{i+NX}^{(n)} \cdot a_{i+NX} - \\ - x_{i+NXY}^{(n)} \cdot a_{i+NXY}) \end{aligned} \quad (8)$$

in which $i = 1...NX \cdot NY \cdot NZ$ (NX points in $x-$direction, NY points in $y-$direction and NZ points in $z-$direction).

The advantage of this method is given by the small amount of memory requirement, precondition work and relatively fast convergency,as well. Through an adaptive determination algorithm for the optimum relaxation factor ω we use only a moderate amount of CPU time [5]. The system of the coupled nonlinear difference equations are solved with Gummel's iterative method.

3 The Numerical Results and Discussion

With the previously given physical model a three–dimensional MOSFET simulation program has been developed. We have investigated several MOSFETs with this program, two of the investigated devices are presented in comparison and the results discussed.

Both investigated devices are of the same geometrical shape and dimensions (Fig. 1 and Fig. 2) except the channel lengths which are $5\mu m$ and $1\mu m$ for of device 1 and 2,respectively. The channel widths are $1\mu m$,the gate oxide thickness $15nm$, the substrate doping $2 \cdot 10^{16} cm^{-3}$ and the source/drain doping $1.69 \cdot 10^{20} cm^{-3}$. In Fig. 1 and Fig. 2 the field oxide which limites the channel in the third dimension, can be seen at the backside of the MOS model. The contacts of source and drain which are left and right in the figures 1 and 2, extend over the channel width, whereas the gate contact covers the channel and the field oxide. The shape of the field oxide in our case is approximated by a rectangular geometry.

The potential distribution at the bias condition $U_{DS} = 2.0V$, $U_{BS} = 0.0V$ and $U_{GS} = 3.0V$ can be seen in Fig. 3 and Fig. 4. The threshold voltages

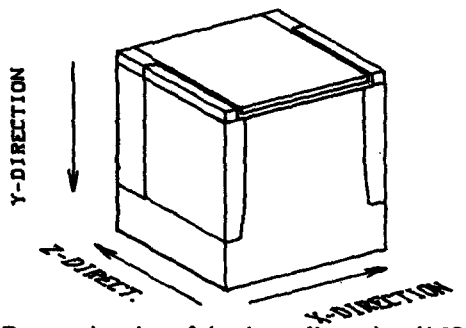


Fig.1: Perspective view of the three–dimensional MOS-FET structure with channel length of $5\mu m$ and channel width of $1\mu m$.

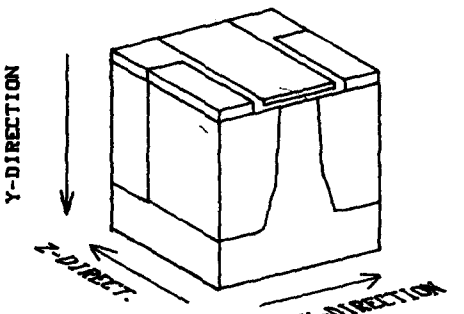


Fig.2: Perspective view of the three–dimensional MOS-FET structure with channel length of $1\mu m$ and channel width of $1\mu m$.

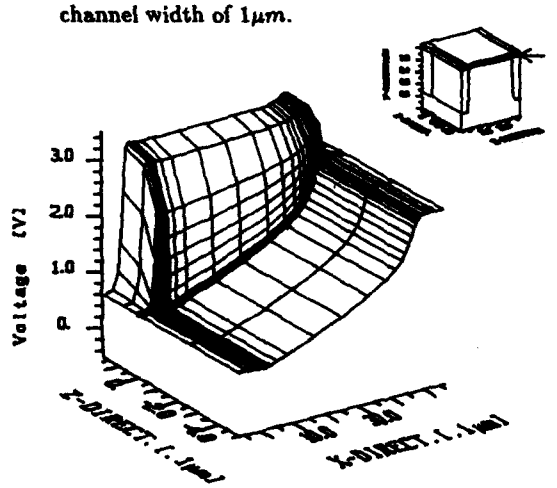


Fig.3: 3D–plot showing a detailed view of the surface potential at the channel edge along the channel length for device 1 at bias $U_{DS} = 2.0V, U_{BS} = 0.0V, U_{GS} = 3.0V$.

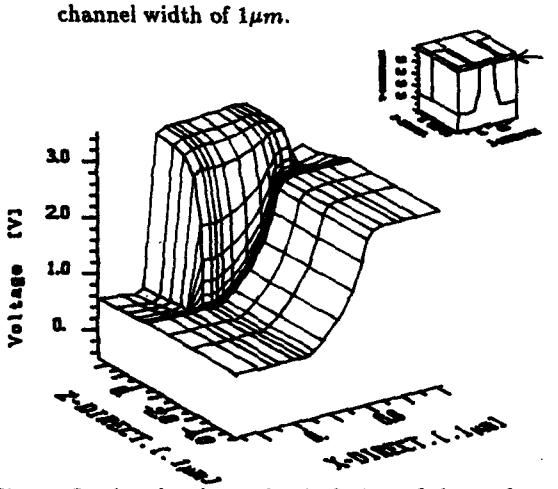


Fig.4: 3D–plot showing a detailed view of the surface potential at the channel edge along the channel length for device 2 at bias $U_{DS} = 2.0V, U_{BS} = 0.0V, U_{GS} = 3.0V$.

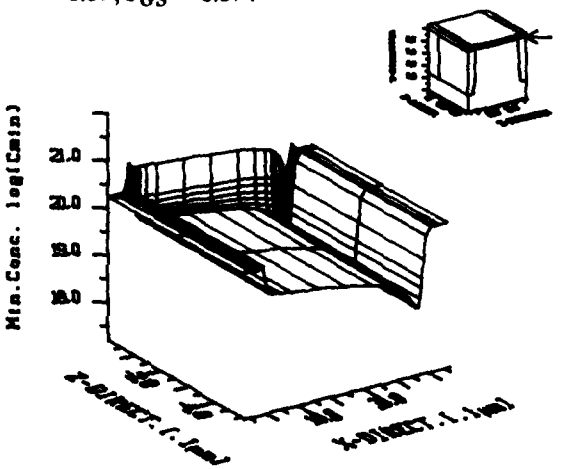


Fig.5: 3D–plot showing a detailed view of the minority density at the channel edge along the channel length for device 1 at bias $U_{DS} = 2.0V, U_{BS} = 0.0V, U_{GS} = 3.0V$.

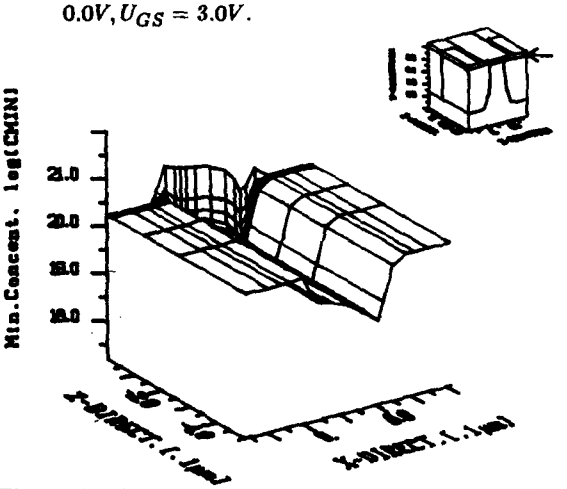


Fig.6: 3D–plot showing a detailed view of the minority density at the channel edge along the channel length for device 2 at bias $U_{DS} = 2.0V, U_{BS} = 0.0V, U_{GS} = 3.0V$.

are $U_{th} = 0.74V$ and $U_{th} = 0.62V$ for device 1 and 2, respectively. That means that we are far above threshold. Note the strong increase of the potential in the field oxide. The minority (electron) distributions (Fig. 5 and Fig. 6), show a very interesting effect. The carrier densities in the short channel MOSFET (device 2) is much higher compared to that of the long channel MOSFET (device 1). The accumulation of the minorities at the channel edge at the given bias condition is based on the limitation of the channel width and the high potential in the oxide region. In the subthreshold region this effect will change into its opposite. The currents calculated by two–dimensional simulations at the previously described bias conditions will be much smaller compared to that of three–dimensional simulations. These effects increase with shrinking device dimensions. Both short and small channel devices will be very sensitive on geometrical changes in respect to their electrical characteristics. Especially we expect that the increase of minorities at the channel edge influences the reliability and safety of devices.

Not only the previously discussed effects but also other three–dimensional effects like the shift of the threshold voltage and the increased breakdown voltage at small channel MOSFETs have been simulated with our program. All these effects are known from theory and from practical measurement but not satisfactorily modeled until now.

Acknowledgement

This work was supported by the research laboratories of SIEMENS AG at Munich, FRG, by DIGITAL EQUIPMENT CORP. at Hudson, USA, and by the "Fonds zur Förderung der wissenschaftlichen Forschung", project S43/10. We are indepted to Prof. H. Pötzl for many helpful discussions.

References:

[1] S.M. Sze, Physics of semiconductor devices, ISBN 0-471-09837-X, John Wiley & sons, 1981

[2] S.M. Sze, VLSI–Technology, ISBN 0-07-062686-3, McGraw-Hill, 1983

[3] T. Iizuka,K.Y. Chiu, and J.L. Moll, Double threshold MOSFETs in bird's–beak free structures, IEEE Int. Electron Device Meet.,Wash.,D.C., 1981, p. 380

[4] L.A. Hageman,Franklin T. Luk,David M. Young, On the equivalence of certain iterative acceleration methods, SIAM J.NUMER.ANAL., pp 852-873, vol. 17 No. 6, Dec 1980

[5] R.G. Grimes,D.R. Kincaid,D.M. Young, ITPACK 2A – A fortran implementation of adaptive accelerated iterative methods for solving large sparse linear systems, Report CNA-164, Center for numerical analysis, University of texas at Austin, 1980

[6] S. Selberherr, The status of MINIMOS, Proc. Simulation of semiconductor devices and processes, pp 2-15, Swansea, 1986

[7] S. Selberherr, Analysis and simulation of semiconductor devices , ISBN 3-211-81800-6, Springer, WIEN NEW-YORK, 1984

[8] O. Axelsson, Solution of linear systems of equations;Iterative methods, Lecture notes in mathematics 574, SMT, 1976

THREE-DIMENSIONAL SIMULATION OF A NARROW-WIDTH MOSFET

P. Ciampolini, A. Gnudi, R. Guerrieri, M. Rudan and G. Baccarani

Dipartimento di Elettronica, Università di Bologna
viale Risorgimento 2, 40136 Bologna, Italy

Abstract

In this paper we illustrate the main features of a general-purpose three-dimensional device-analysis program, HFIELDS-3D, developed at the University of Bologna in the context of an EEC-supported ESPRIT Project. The program employs triangular-based prismatic elements, which provide a reasonable compromise between simplicity and flexibility, but it is not otherwise limited to any specific device structure. In the present implementation, the program handles Poisson and one-carrier continuity equation, which allows for the simulation of unipolar devices. As an example, a typical 3-D problem, the narrow-width effect, is investigated using a realistic device structure fully accounting for the typical bird's beak. It is shown that not only the threshold voltage, but also the gain factor, and therefore the device transconductance, are affected by the narrow-width effect.

1. Introduction

Numerical simulation of semiconductor devices in two dimensions is nowadays a well-established technique for the design of advanced electronic components and processes. As device miniaturization progresses toward submicron feature sizes, however, 3-D effects are getting more and more important even for nominally-standard planar devices, thus making two-dimensional simulation codes inadequate for predicting device performance. In addition, increasingly complex device geometries are being devised, such as the buried-electrode dynamic RAM cell, currently used in high-capacity memory devices, the floating-gate EPROM cell, and the I^2L NOR gate, which are inherently three dimensional. All the above devices can only be simulated by means of 3-D device-analysis programs.

Most of the activity reported so far in this field has been performed in Japan [1-3] and, to a lesser extent, in the United States [4]. In this paper we illustrate the main features of a general-purpose three-dimensional code, HFIELDS-3D, developed at the University of Bologna in the context of an EEC-supported ESPRIT Project, and show how such a code can be profitably used to investigate a classical three-dimensional problem, i.e. the narrow-width effect in MOSFET's.

The program employs triangular-based prismatic elements, which allow for a reasonable compromise between geometrical flexibility and simplicity of implementation, but it is by no means restricted to any specific device structure. In the present stage of development, only Poisson and one carrier-continuity equation are solved, thus making the program suitable for the simulation of unipolar devices.

The next section discusses the fundamental choices of the present project, and provides some details on the software implementation of the program. Numerical results are illustrated in section 3 and conclusions are drawn in section 4.

2. Features of the program

HFIELDS-3D allows up to ten semiconductor and insulator regions, which can be either simply or multiply connected. An equal number of insulator-semiconductor interfaces, ohmic contacts, gates and floating gates are allowed. Thus, rather complex device structures can be accommodated by the program.

As already anticipated, HFIELDS-3D employs a triangular-based prismatic-element mesh which, in the authors' opinion, is flexible enough from the geometrical standpoint for most practical applications, while still having a number of implementation advantages over tetrahedral meshes:

- The problem of properly defining the control volumes associated with the nodes of general tetrahedra is still largely unsolved or, at least, insufficiently tested. As a result, some ripple could be expected in the resulting solution.
- The problem of generating tetrahedral meshes in three dimensions which take real advantage of the potential flexibility of these elements (i.e. without converting prisms into tetrahedra) is a very hard task. On the other hand, generating a prismatic-element mesh can be easily accomplished by simply replicating a triangular mesh in the third dimension. It should be noticed that the above procedure does not imply a geometrical uniformity of the simulated device in the third dimension, but it requires step-like changes.

In its present implementation the program handles Poisson and one carrier continuity equation (electrons), which are solved using the Gummel [5] successive procedure. The adopted discretization scheme is the well-known "box integration method", whereby each node is allocated a control volume which, in our case, is still a prism. Poisson's equation is discretized assuming a piecewise linear approximation for the electric potential along the mesh lines, and a three-dimensional generalization of the Sharfetter-Gummel scheme [6] is used to discretize the current-continuity equations. The discretization procedure of Poisson's equation leads to a set of N non-linear algebraic equations, while the corrent-continuity equation leads to a linear system (if the dependence of the carrier mobility upon carrier concentration is neglected).

For the solution of the linear system, we employ the ICCG method for Poisson's equation, and a direct solver for the current-continuity equation. Parallel techniques are currently being investigated for a subsequent vector-processor implementation of the code.

3. The narrow-width MOSFET

Several authors [7,8] have published experimental and/or theoretical results on narrow-width MOSFET's. From the above papers, one can infer that the most important narrow-width effect is an increase of the threshold voltage as the channel width becomes narrower.

We have simulated the electrical properties of a narrow-width MOSFET having a fixed, nominal channel length $L = 1.0\,\mu m$ and a nominal channel width W ranging from $0.4\,\mu m$ to $2.8\,\mu m$. The mesh of one of the above devices is shown in fig. 1: due to the simmetry of the structure, only half of the device is actually considered. The "bird's beak" at the transition between the channel and the field regions is carefully described on the front plane in order to take full advantage of the flexibility of the triangular mesh; so doing, current flow occurs mainly in the third direction, i.e. normal to the front plane.

The 2-D mesh is replicated in the third dimension, where 19 planes are accommodated, with slight modifications which account for the structural changes of the MOSFET (transitions between the source-gate and gate-drain regions). The whole mesh comprises 2,565 nodes and 4,158 prisms.

The impurity concentration is input via analytical expressions reasonably accounting for source and drain diffusions, channel implant and channel-stop diffusion. The channel implant was designed to give a threshold voltage $V_T = 0.7\,V$.

All simulations were carried out on a MicroVAX-GPX work station, and the average CPU time required was about 1.5 hours per bias point.

Figure 2 shows a perspective plot of the equilibrium potential in the plane normal to the X-axis, located at the field-oxide silicon interface. The two upper "plateau" represent the source and drain regions, and the channel appears as a saddle between them. Figure 3 shows instead the electric potential in the parallel plane located at the gate-oxide silicon interface. The ridge at the periphery of the channel is due to the penetration of the device cross section into the oxide and not to an actual increase of the electric potential at the Si-SiO_2 interface. Rather, the fringing field arising from the field-oxide penetrates in the channel region, causing a decrease of the electric potential at its edges and thus reducing the effective inversion layer width. Consequently, when applying a drain-source voltage, a corresponding decrease in the corrent flowing along the channel is to be found.

In order to highlight this effect, some comparisons with 2D results have been made: a corresponding planar device has been simulated neglecting the field oxide and the bird's beak, and assuming equal channel widths. Figure 4 shows the simulated turn-on characteristics of the 2-D and 3-D MOSFET's for three different values of the channel width, namely $W = 0.4\,\mu m$, $W = 1.2\,\mu m$ and $W = 2.8\,\mu m$, respectively. The drain-source voltage $V_{DS} = 0.1\,V$ in these simulations. The figure shows that the 2-D MOSFET systematically provides a larger current than the 3-D one, and the difference increases as the gate voltage is increased. Thus, both threshold voltage and gain factor are affected by the narrow-width effect. Figure 5 shows the corresponding turn-on characteristics with $V_{DS} = 3.0\,V$, which confirm the above statement. The subthreshold behaviour of the turn-on characteristics is better illustrated in figure 6. For the 0.4 μm-width device, over one order of magnitude difference between the 2-D and 3-D models is observed.

A comprehensive view of the I_D/W dependence upon gate voltage and channel width is given in figures 7 and 8, where a perspective plot of a 2-D surface in the 3-D space is shown. These plots emphasize the transconductance degradation and the change in threshold voltage which occur for small values of the channel width.

4. Conclusions

In this paper we have illustrated the main features of a general-purpose three-dimensional device-analysis program, HFIELDS-3D, developed at the University of Bologna in the context of an EEC-supported ESPRIT Project. Care has been taken to ensure versatility and geometrical flexibility of the code, which allows for a wide variety of realistic device structures. At the present stage of development, the program solves only Poisson's and one carrier-continuity equation, which makes it suitable for the simulation of unipolar devices, but its extension to both carrier-continuity equa-

tions is forseen in the near future. The program has been shown to perform satisfactorily in the investigation of the narrow-width effect in MOSFET's, but much work is still to be done in order to optimize its numerical efficiency. To this purpose, the authors are currently investigating more efficient algorithms for a vector-processor version of the code. In order to handle the huge number of equations inherent in 3-D problems in a reasonable time, the use of vector processors or parallel architectures is mandatory.

Acknowledgements

This work has been partially supported by EEC under the ESPRIT 962E-17 Project.

References

1. T. Toyabe, H. Masuda, Y. Aoki, H. Shukuri and T. Hagiwara: "Three-dimensional device simulator CADDETH with highly convergent matrix solution algorithms", *IEEE Trans. on Electron Devices*, vol. ED-32, pp. 2038-2043, 1985.
2. K. Yokoyama, M. Tomizawa, A. Yoshii and T. Sudo: "Semiconductor Device Simulation at NTT" *IEEE Trans. on Electron Devices*, vol. ED-32, pp. 2008-2017, 1985.
3. N. Shigyo and R. Dang: "Three-Dimensional Device Simulation Using a Mixed Process/Device Simulator", from *Process and Device Modeling*, Ed.: W. L. Engl, North Holland, 1986.
4. E. M. Buturla, P. E. Cottrell, B. M. Grossman, C. T. McMullen and K. A. Salsburg: "Three-Dimensional Transient Finite-Element Analysis of the Semiconductor Transport Equations", from *Numerical Analysis of Semiconductor Devices, Proc. of the NASECODE II Conference*, pp. 160-165, Boole Press, Dublin, 1981.
5. H. K. Gummel: "A Self Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculation", *IEEE Trans. on Electron Dev.*, vol. ED-11, pp. 455-465, 1964.
6. D. L. Sharfetter and H. K. Gummel: "Large-Signal Analysis of a Silicon Read Diode Oscillator", *IEEE trans. on Electron Devices*, vol. ED-16, pp. 64-77, 1969.
7. L. A. Akers, M. M. E. Beguwala and F. Z. Custode: "A Model for a Narrow Width MOSFET Including Tapered Oxide and Doping Encroachment", *IEEE Trans. on Electron Dev.*, vol. ED-28, pp. 1490-1495, 1981.
8. Y. C. Cheng and P. T. Lai: "An Analytical Model for the Threshold Voltage of a Narrow-Width MOSFET", *IEEE Trans. on Electron Dev.*, vol. ED-31, pp. 1814-1823, 1984.

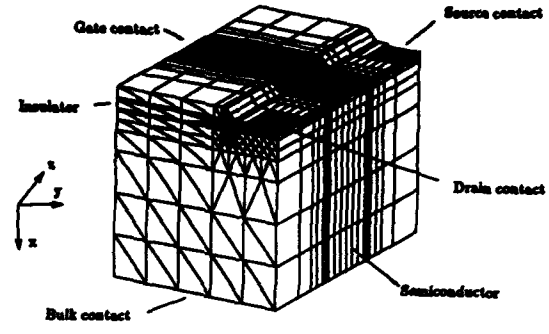


Fig. 1: Mesh of the simulated MOSFET.

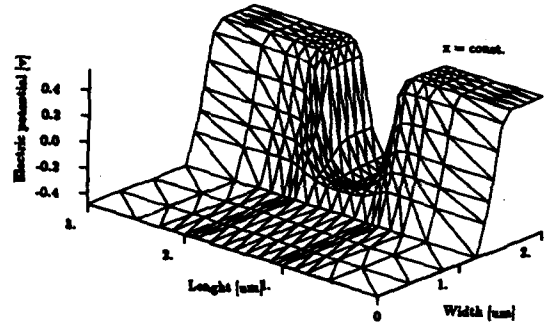


Fig. 2: Perspective plot of the equilibrium potential in the plane normal to the X-axis, located at the field-oxide silicon interface.

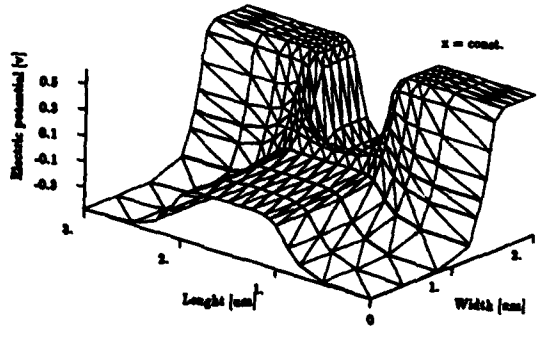


Fig. 3: Perspective plot of the equilibrium potential in the plane normal to the X-axis, located at the gate-oxide silicon interface.

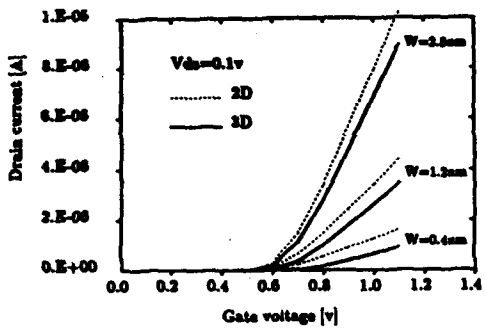


Fig. 4: Simulated turn-on characteristics of the 2-D and 3-D MOSFET's for different values of the channel width; $V_{ds} = 0.1V$.

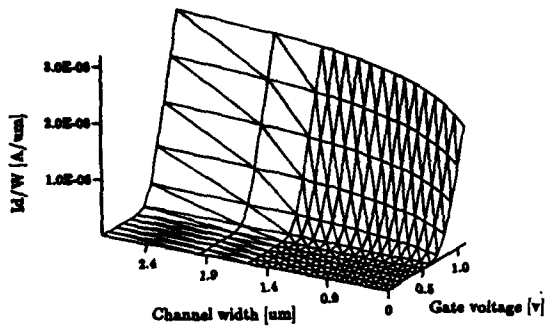


Fig. 7: Normalized drain current as a function of the channel width and the gate voltage.

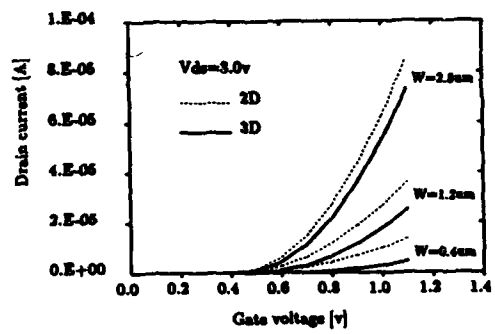


Fig. 5: Simulated turn-on characteristics of the 2-D and 3-D MOSFET's; $V_{ds} = 3.0V$.

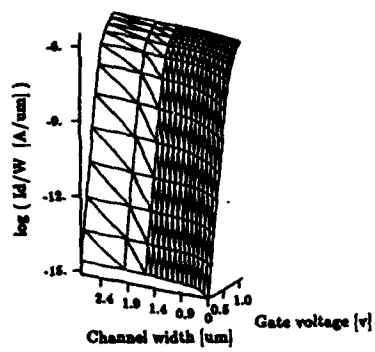


Fig. 8: Normalized drain current as a function of the channel width and the gate voltage.

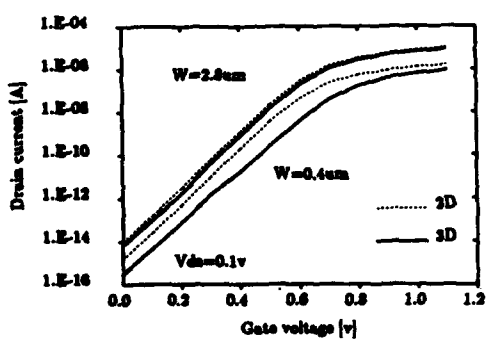


Fig. 6: Simulated turn-on characteristics of the 2-D and 3-D MOSFET's; $V_{ds} = 0.1V$.

2-D AND 3-D CAPACITANCE EFFECTS IN MOS VLSI

J.H.M.M. Quint, F.M. Klaassen, R. Petterson

Philips Research Laboratories
P.O.Box 80000, 5600 JA Eindhoven, The Netherlands

ABSTRACT

Spreading capacitances of several MOS VLSI configurations have been calculated numerically by solving Poisson's equation in 2 or 3 dimensions. Owing to nonuniform charge distributions, contributions from sidewalls and topsurfaces, and shielding effects, considerable deviations from scarce analytic formula have been found. Successively considered are the cases: 3 parallel conductors at equal height from the substrate, 2 parallel conductors at different level from the substrate, gate-drain configuration of different MOSFETs, and two conductors or four conductors crossing above a substrate.

1. Introduction

Since bias voltage constraints and electromigration effects prevent to scale down properly the dimensions of the interconnection system, in submicron VLSI additional capacitance effects become more and more important. Not only do charges at the sidewalls and top surfaces of the conductors lead to a larger capacitance to substrate or to interline capacitance, but nonuniform charge distribution and shielding effects cause the capacitance to deviate considerably from the 1-D value.

Although analytic capacitance expressions based on conformal mapping have been given to correct for several of the above effects [1,2,3], it is questionable whether these results are generally useful. Not only are the results limited owing to several assumptions used, but in practical layouts configurations soon become too complicated to use conformal mapping. Since an empirical investigation is costly and sometimes even impossible, the capacitances from a number of elementary interconnect configurations have been calculated numerically by making use of the device simulator TRIPOS [4], which solves Poisson's equation in two and three dimensions. Where possible, the results have been compared to analytic results. Furthermore the observed trends are discussed from a physical point of view.

2. Capacitances between 3 parallel conductors and the substrate

Owing to 2-D and 3-D charge distribution effects, the capacitance per unit length deviates from the conventional 1-D formula $C = \epsilon_{ox} W/H$, when the distance (H) to substrate and the thickness (T) of conductors are no longer small compared to their width (W) and mutual distance (S). This is shown in fig. 1, where all possible capacitances of 3 conductors parallel to each other and to the substrate are given as a function of the distance S (with $T = W = 1$ μm). Lateral field distributions underlying the above 2-D results are given in fig. 2 (with the location indicated in the inset of fig. 1). While the normal field remains rather uniform under the central conductor owing

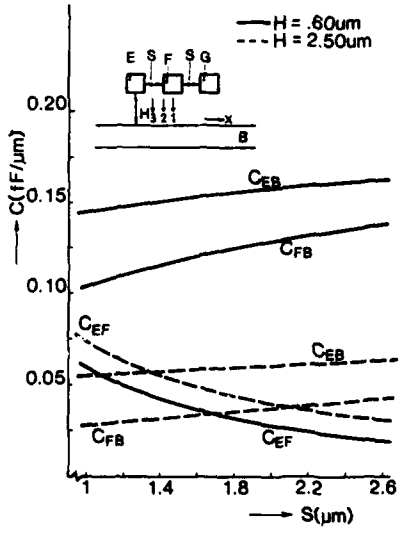


Figure 1. Capacitances vs. mutual distance for conductor configuration given in the inset.

to shielding by the surrounding conductors, the lateral field is very nonuniform. Starting from nearly zero value along the symmetry line (1), the above field peaks to a high value at the edge (2) of the central conductor due to charge crowding at the corners of the conductor (at which upto 30% of the total charge may be accumulated).

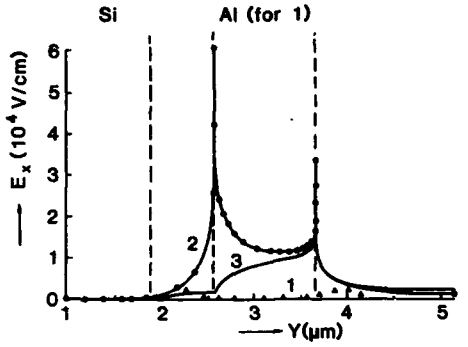


Figure 2. Lateral field along several normal directions for configuration of fig. 1.

Owing to shielding effects of field lines originating from the sidewalls and top surfaces the capacitance to substrate of the central conductor ($C_{FB} = \Delta Q_B/\Delta V_F$) is lower than that of the left conductor and for smaller values of the distance S the difference becomes larger. At the same time the interline capacitance ($C_{FE} = \Delta Q_E/\Delta V_F$) increases with decreasing S and at a higher rate, than the others decrease. As shown in fig. 3 the total capacitance of the middle conductor $C_T = C_{FB} + 2C_{FE}$ increases with decreasing S. In the same figure also a semi-empirical approximation [3] is given. Although the deviation with TRIPOS remains within 10%, the deviation for the various parts of C_T is larger.

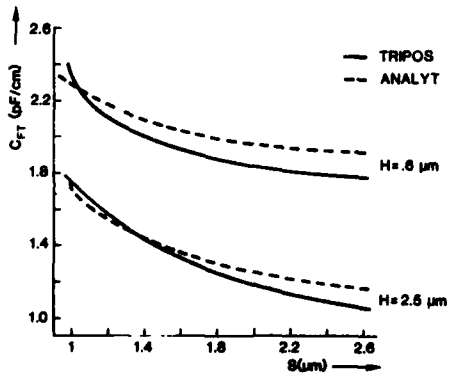


Figure 3. Total capacitance of central conductor of fig. 1.

3. Capacitances between two conductors at different level

When two conductors are parallel, but located at different levels compared to the substrate, their partial overlap has a specific effect on all possible capacitance values. This is shown in fig. 4, which gives the capacitances as a function of the overlap distance S as defined in the inset. Owing to shielding of field lines originating from the top surface of conductor L, the capacitance to substrate C_{LB} at complete overlap ($S = 0$) is lower than the value in the absence of conductor T, and even slightly decreases at small partial overlap ($S < 0.3$ μm). Only when the overlap approaches zero ($S > 1.0$ μm) this capacitance increases with S. Because of differences in the shielding of field lines originating from the bottom and sidewalls of conductor T, for the capacitance C_{TB} the same qualitative behaviour is observed. However, owing to the fact that at $S = 0$ the bottom surface of T has been shielded effectively, the relative effects are smaller. Furthermore, while at complete overlap the interline capacitance C_{TL} varies strongly with the step height ($H_T - H_L$), for larger values of S the above capacitance becomes completely determined by fringing effects and therefore hardly varies.

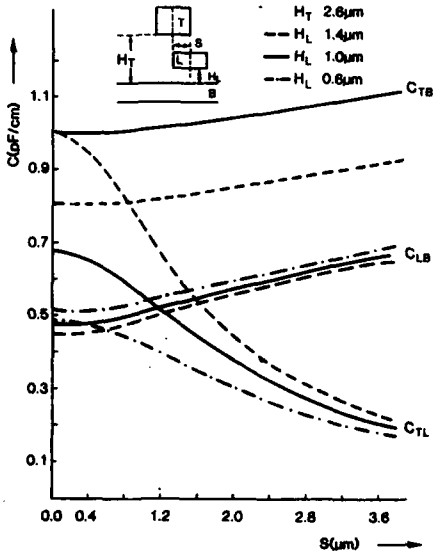


Figure 4. Capacitances vs. mutual distance for configuration given as inset (thickness of L = .6 μm).

4. Gate-drain capacitance of a MOSFET

Below threshold voltage the gate-drain capacitance ($C_{GD} = \Delta Q_G/\Delta V_D$) is larger than expected from the physical overlap owing to fringing effects between the poly silicon gate sidewall and the drain topsurface, and

between the drain-channel edge and the gate lower surface (compare the inset of fig. 5). For the conventional MOSFET an analytic expression for the above capacitance

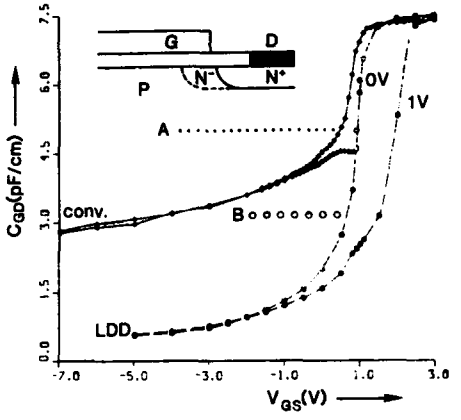


Figure 5. Gate-drain fringing capacitance of conventional and LDD MOSFET (lines (A, B) are analytic results).

has been given [5], but the validity of this approach is questionable for an LDD-MOSFET with its refined gradual drain junction profile. For both devices the numerically calculated value of C_{GD} is given in fig. 5 as a function of the gate voltage at two different values of the drain voltage. The devices considered are n-channel type with a poly silicon gate (width 0.7 μm, thickness 0.3 μm) on a 17 nm thin insulator. For the conventional MOSFET the 0.25 μm drain junction has a physical overlap with the gate of 0.17 μm. Using process modelling data the lightly doped drain of the LDD-type is assumed to overlap the gate to an amount of 0.08 μm.

Decreasing V_G from zero Volt (off-state) causes the building-up of an accumulation charge starting from the centre of the channel region. This growing accumulation charge gradually reduces the inner fringing capacitance between gate and drain edge. In the sub-threshold region C_{GD} increases with V_G owing to the gradual formation of a channel starting from the drain. This generally increases the 1-D overlap. At threshold voltage the capacitance saturates with V_G, since the channel is present everywhere. However, due to channel pinch-off occuring at V_D = 1 Volt C_{GD} remains lower than the value at V_D = 0. Owing to the smaller 1-D overlap and partial depletion of the lightly doped drain region (in particular at V_D = 1 Volt) the capacitance of the LDD MOSFET is considerably smaller in the accumulation and subthreshold region.

Also shown in fig. 5 is an analytic result [5]. Only for the conventional device at V_G = 0 this result fairly agrees with the numerical value. However the reduction of the fringing effect in deep accumulation is not taken into account.

5. Capacitance of two crossing conductors

When a second conductor crosses a first one at different height from the substrate (inset of fig. 6) the field and the resulting charges have to be calculated using a full 3-D solution of the Laplace equation [6]. Fig. 6 gives the three possible capacitance values as a function of distance between both conductors. The upper conductor T and the lower conductor S have a thickness of 1.1 μm and 0.6 μm respectively, a width of 1 μm and a length of 20 μm.

Of course the most pronounced effects are found in the interline capacitance C_{TS}. Not only is the value owing to 3-D charge contributions almost an order higher than the 1-D capacitance value, but the presence of the substrate causes that the above value is still smaller than the pure interline capacitance value C'_{TS} (substrate removed). In addition the contributions from sidewalls and top/bottom areas cause the capacitance C_{TS} to vary sublinearly with H^{-1}. Shielding effects are the main reason for C_{SB} to increase slightly with increasing H.

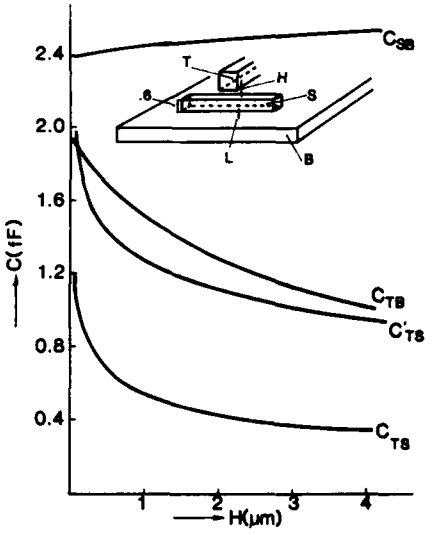


Figure 6. Capacitances vs. distance for two crossing conductors above a substrate (L = 1 μm).

6. Capacitances of four conductors

As an extension to the previous case a configuration of four conductors is considered, which are crossing mutually and are located at two different levels above the

substrate (compare the inset of fig. 7). Naturally the same 3-D effects occur as in the previous case, but owing to the more complicated situation the calculation is more demanding for the simulator. In spite of making use of symmetry rules the number of grid points required amounts 30.000.

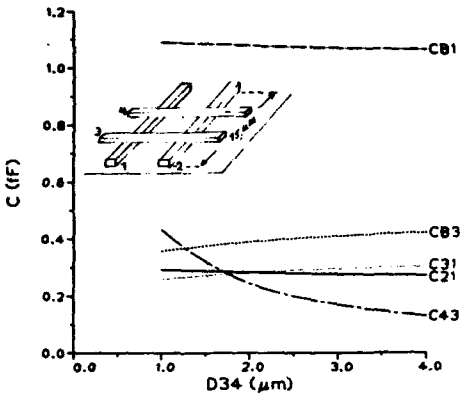


Figure 7. Capacitances between four conductors (see inset) vs. distance of upper conductors ($d_{13} = d_{1B} = 1\ \mu m$, $d_{12} = 2\ \mu m$).

In fig. 7 several capacitances to bulk and interline capacitances have been plotted as a function of the distance between the two upper conductors. Except the capacitance C_{34} the variation of the other capacitors is dominated by shielding effects. C_{B1} and C_{21} decrease slightly with increasing distance d_{34} owing to the fact that in this case the upper conductors have a more effective shielding on the lower conductors. On the other hand with increasing distance d_{34}, conductor 4 causes less shielding of field lines originating from conductor 3. Therefore C_{31} increases with in creasing d_{34}. Although the lower conductors and the substrate have a shielding effect on C_{34}, this capacitance still varies with d_{34}^{-1}.

Finally in fig. 8 the insulator thickness d_{13} between the upper and lower conductors has been varied. Owing to a reduction of shielding effects the interline capacitances between conductors at the same level increase considerably with an increase of d_{13}. Naturally the capacitances between conductors at different levels decrease at the same time, but similar to the case of fig. 6 this decrease is only a sublinear function of the thickness.

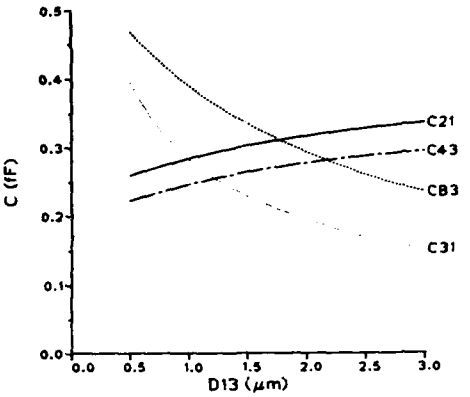


Figure 8. Capacitances vs. insulator thickness ($d_{12} = d_{34} = 2\ \mu m$, $d_{1B} = 1\ \mu m$).

Conclusions

When comparing the results to the limited number of analytic expressions (case 1 and 3), already strong deviations are observed. All trends or effects observed in the figures can be explained from a physical point of view. A change of interconnect dimensions or mutual distances has much larger effect than the thickness of dielectric layers. For 1 μm wide conductors 2-D capacitance values may exceed the 1-D value by a factor of 5, and in 3-D cases by a factor of 9. Estimating 2-D capacitance values by reducing surface charge integrals to line integrals or 3-D capacitance values as a summation of 2-D cases is too inaccurate.

References

[1] M. Elmasry, IEEE **EDL-3**, 1 (Jan. 1982).
[2] T. Sakurai et al., IEEE **ED-30**, 2 (Febr. 1983).
[3] E.W. Greeneich, IEEE **ED-30**, 12 (Dec. 1983).
[4] D.J. Coe et al., Philips proprietary 3-D device simulator.
[5] R. Shrivastava, IEEE **EDL-6**, 3 (March 1985).
[6] P.E. Cottrell et al., IBM J. **R/D-29**, 3 (May 1985).

Session P2.1

Posters

Tuesday, September 15, 1987

IMPLANTATION and DIFFUSION MODELLING of BORON in SILICON

An De Keersgieter, Luc Dupas, Kristin De Meyer*
IMEC, Kapeldreef 75, B-3030 Leuven, Heverlee, Belgium

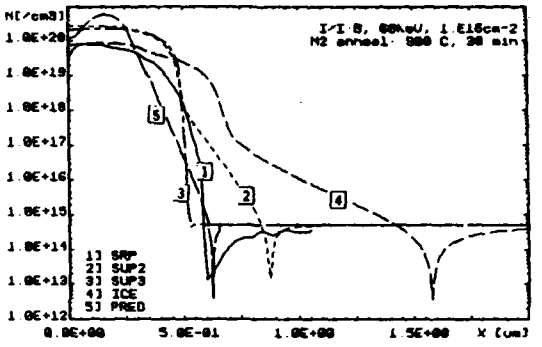
Abstract

The accurate simulation of implanted and diffused impurity profiles in silicon is extremely important when developing VLSI processes. In this work simulations with different process simulators and the corresponding experimental results for implantation and diffusion in N_2 ambient of boron in silicon are compared. Our study reports a remarkable dose dependence of the shape of the experimental profiles. Strategies have been developed to increase the simulation accuracy.

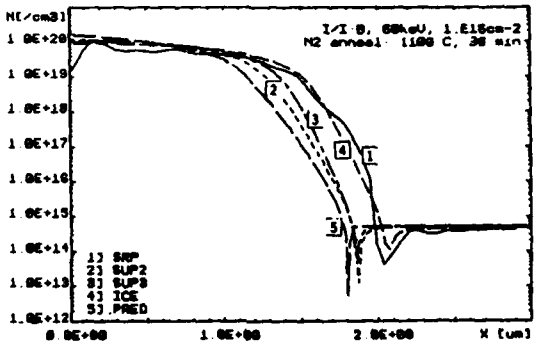
1 INTRODUCTION

The simulators used in this study are: SUPREMII.5[1], SUPREM3[2], ICECREM[3] and PREDICT[4]. The substrates used were 9-13 Ωcm phosphorus-doped n-type (100) silicon wafers. A 400 $\AA$ thick thermal oxide was grown in dry oxygen at 1000°C. Then 60 keV, $^{11}B^+$ implants with doses ranging from 10^{12} cm^{-2} to $10^{17}cm^{-2}$ were performed with wafers tilted 7° to suppress channeling. The annealing temperatures and times were : 900, 1000, 1100 °C and 30, 300, 1200 min.

Comparing the experimental implantation/diffusion profiles and simulations (using default model parameters) no satisfactory agreement is achieved. It is found that for long, high temperature anneal steps simulations and experiments coincide far much better than for low temperature, short time annealing conditions (Figure 1). In the latter case the actual shape of the implantation (starting) profile has a pronounced influence on the final diffusion profile. This is illustated in Figure 2 for profiles simulated with ICECREM. For a 30 min anneal at 1000°C we still see a tail in the profile which is originating from a strong implantation channeling model. After 300 min of annealing the shape of the profile has completely changed. So, the exact modelling of the implantation profile is not only necessary for improving the accuracy of the final diffused profiles, but it is also very important for a better understanding and comparison of the diffusion mechanisms . Also modelling of rapid thermal annealing has a need for precise implantation start profiles.



a)



b)

FIGURE 1 :

Comparison between simulations and experiments
a) for low temperature : 900°C, 30 min
b) for high temperature : 1100°C, 30 min

*Professor of the Katholieke Universiteit, Leuven, Research Associate of the Belgian National Found for Scientific Research.

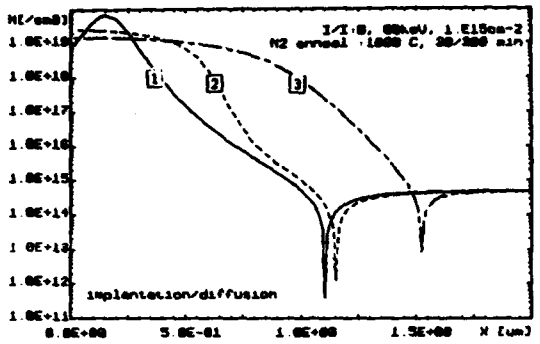


FIGURE 2 :

ICECREM default profiles :
1 as implanted profile : B, 60keV, 10^{15} cm^{-2} (400 Å oxide)
2 same profile after N_2anneal, 1000°C, 30 min
3 same profile after N_2anneal, 1000°C, 300 min

2 DEFAULT IMPLANTATION MODELLING

Most process simulators use analytical formulas for modelling the implanted profiles. Here different models and also different default parameter values are found when comparing the available programs. Most frequently used are the Gauss model (non-channeled profiles), or Gauss-related, PearsonIV and PearsonIV-related (profiles with channeling tail) models. The model chosen depends on the specified impurity, on the type of target, on the thickness of the implant oxide and also on the simulator used (Table 1).

SUPREMII.5	Gauss		PearsonIV + exp.tail
ICECREM	Gauss		PearsonIV
SUPREM3	Gauss	2-Gaussian	PearsonIV
PREDICT			Gauss + exp. tail

Table 1: Different process simulators and their models for implantation of boron in crystalline silicon

The Gauss model

$$C(x) = \frac{Dose}{\sqrt{2\Pi}\sigma} exp\left(-\frac{(x-R_p)^2}{2\sigma^2}\right)$$
$$= Dose \times f(x) \quad , f(x): frequency function$$

gives a symmetrical profile which can only be used for non-channeled implantation profiles (e.g. in amorphous silicon). R_p is the range of the profile and σ is the standard deviation.

In crystalline silicon channeling occurs and the profile is no longer symmetrical. Then at least one has to account for the skewness γ by using a frequency function which consists of two half Gaussian distributions. If also the peakedness β is needed to describe the asymmetry, the PearsonIV distribution is used. This distribution is given by [5]:

$$C(x) = Dose \times f(x)$$
and
$$f(x) = K[-(b_o + b_1 y + b_2 y^2)]^{-1/2b_2} \times exp\left(-\frac{b_1/b_2 + 2a}{\sqrt{4b_o b_2 - b^2}} \times arctan\frac{2b_2 y + b_1}{\sqrt{4b_o b_2 - b^2}}\right)$$

a, b_o, b_1, b_2 are functions of σ, γ, β

In this study we concentrate on boron implantations in crystalline silicon, unannealed or annealed in N_2 atmosphere. As all experimental as-implanted profiles (measured by SIMS) show a certain amount of channeling (even with a 400 Å thick implant oxide and 7° tilt) PearsonIV, modified PearsonIV and modified Gaussian models are used (column 4 of Table 1). Using the default implantion models and parameters none of the process simulators gives a satisfactory agreement with the experimental as-implanted profiles over a wide range of doses (Figure 3). SUPREMII.5 and ICECREM show the same tendencies. In both cases the tail extends beyond the extrapolated tail from the SIMS measurement for doses of 10^{15} cm^{-2} and higher. For 10^{13} and 10^{14} cm^{-2} simulations and experiments agree considerably well and for the lowest dose (10^{12} cm^{-2}) simulations seem to underestimate channeling. PREDICT is the only simulator which accounts for a dose dependency for channeling, but the modelling is done in a empirical and discontinuous way. SUPREM3(version 3C) does not model channeling at all. Also remarkable is that literature about SUPREM3 always mentions a PearsonIV model while the default implantation parameters satisfy the conditions for a Pearsonl distribution.

Our study thus revealed a remarkable dose dependence of the shape of the experimental profiles (especially in the channeling tail region) which is completely neglected by the available simulators (except to some extend for PREDICT). So it is obvious that the default implantation parameters cannot be used when simulating the actual implanted profiles and as such will also provide a non-optimal starting profile for the study of subsequent diffusion steps.

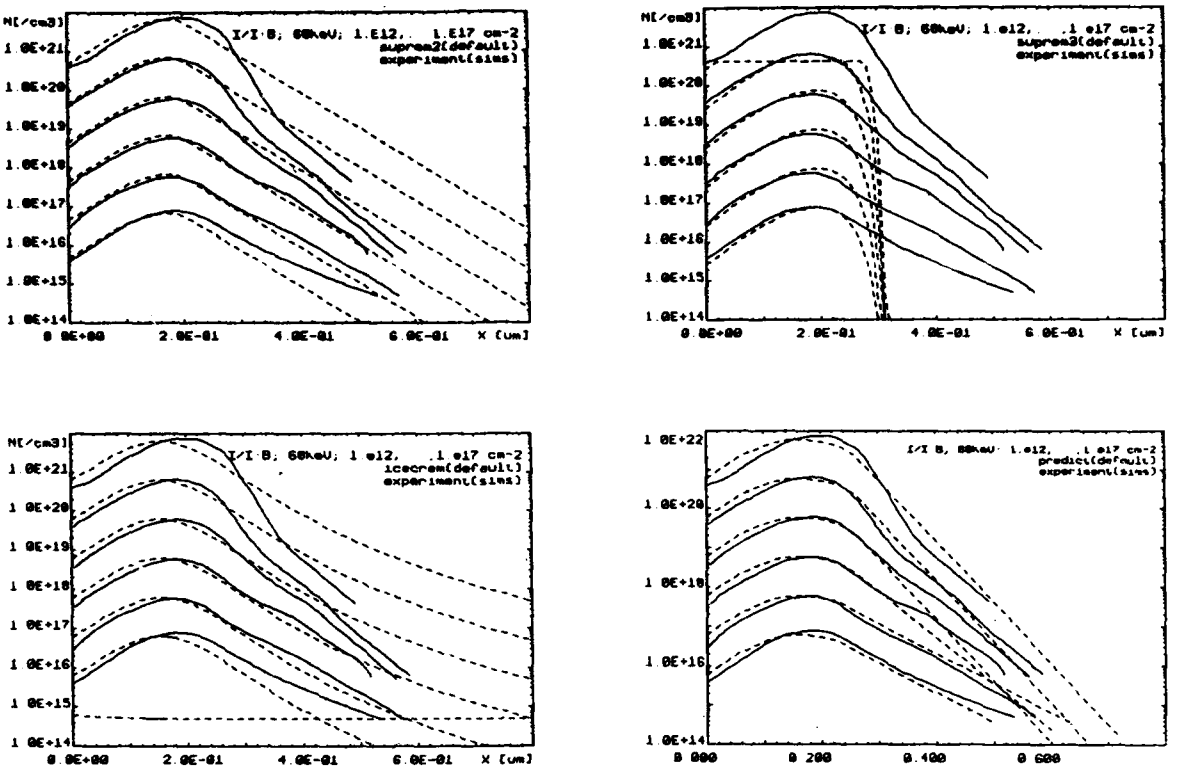


FIGURE 3 :

Comparison between different process simulations and experimental SIMS implantation profiles for I/I: B, 60keV, 10^{12}, 10^{13},...,10^{17} cm^{-2} through 400 $\AA$ oxide. The solid lines represent SIMS-data. The origin of the x-axis is at the Si/SiO_2 interface.

3 NON DEFAULT IMPLANTATION PROFILES

In our work we used the PearsonIV model (used in ICECREM) for extracting accurate model parameters from the experimental SIMS profiles using the SIMPAR[5] package. This parameter extraction program has originally been developed for determining the parameters used in the I-V relations of transistors, but it is also possible to built in any user defined model.

Very good PearsonIV fittings for the experimental implantation results are obtained (Figure 4). The strong channeling effect at higher doses has disappeared, but still some tailing depending on the dose, can be observed. This fitting procedures leaves all distribution parameters ($R_p, \sigma, \gamma, \beta$) free. This is of course not very practical for process simulation, but the resulting sets of parameters do exhibit a dose dependence which can be quite easily described. To model this dependence in a phys-

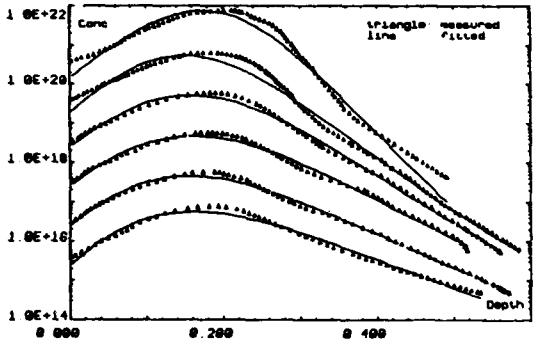


FIGURE 4 :

SIMPAR (PearsonIV) fitting results for implantation

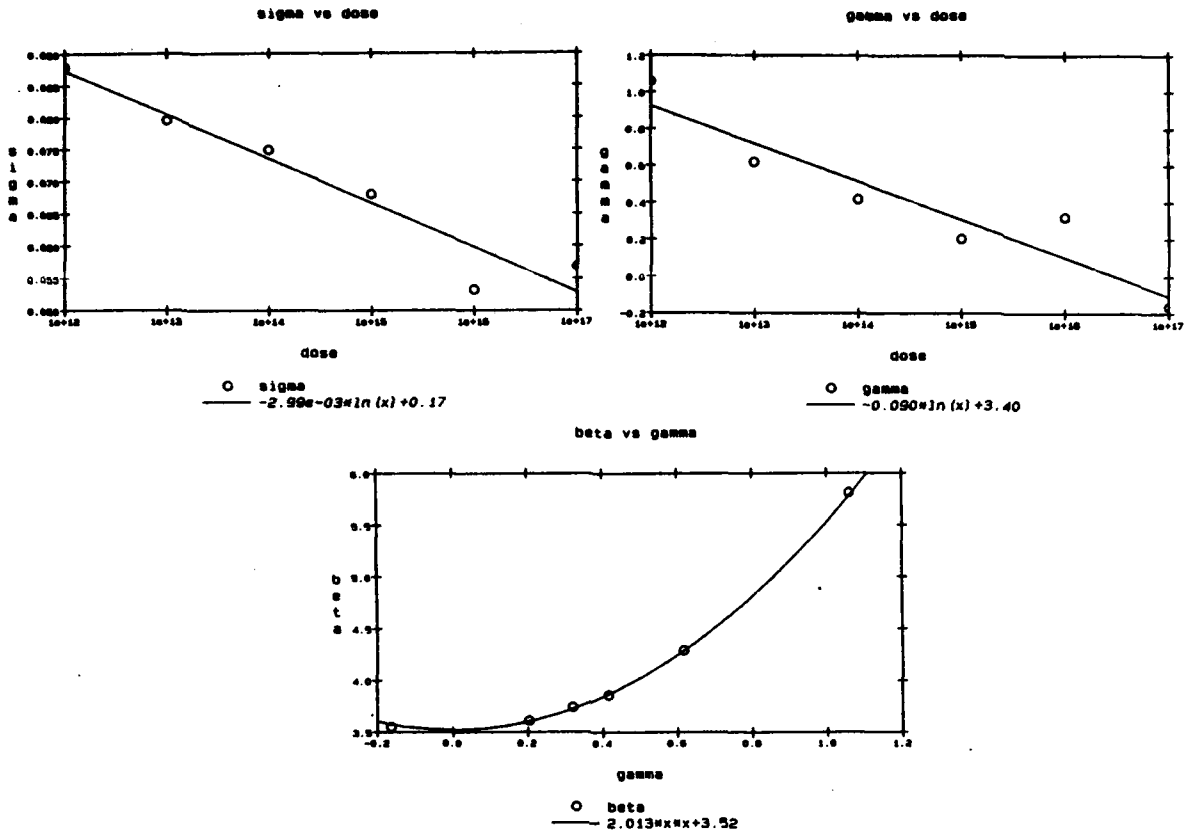


FIGURE 5 :

Dose dependent implantation parameters : σ vs dose, γ vs dose, β vs dose
The circles represent extracted values.

ical and practical way, a second fitting procedure is performed where R_p is kept fixed at a mean value corresponding to the implantation energy. This means we consider R_p independent of the implantation dose. This approach is generally accepted. On the other hand, σ, γ, β are still clearly dose dependent. This explains also the variation of the position of the maximum $= R_p + a$, with a a function of γ and β. The results for a 60keV boron implantation through 400 $\AA$ oxide are given by :

$$\begin{cases} R_p & \approx & cte \\ \sigma & \approx & -2.99 10^{-3} ln(dose) + 0.17 \\ \gamma & \approx & -0.090 ln(dose) + 3.40 \\ \beta & \approx & 2.013\gamma^2 + 3.52 \end{cases}$$

The extracted parameters and the resulting relationsships are shown in Figure 5. All non-constant parameters are decreasing with increasing implantation dose. The skewness γ and the peakedness β are approaching to Gaussian values ($\gamma = 0, \beta = 3$) as the dose increases. This means the higher the dose the less channeling occurs. Indeed for higher doses more damage is created, which reduces the probability for channeling.

4 DIFFUSION PROFILES

Finally starting from the proper implantation profile, different diffusion models can be compared. All diffusion profiles were measured by spreading resistance profiling (SRP). In general we now get a much better agreement with the experimental results, also for the low temperature, short time anneal conditions. Indeed, the strong channeling tail, which could be observed for default simulations and is not at all occuring in the experimental profiles, is eliminated (Figure 6, Figure 7).

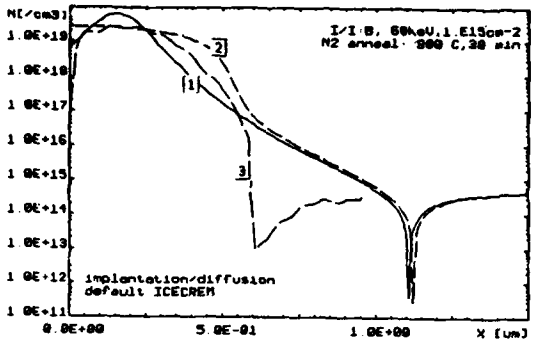


FIGURE 6 :

ICECREM default profiles vs experimental SRP profile :
[1] ICECREM implantation B, 60keV, $10^{14}cm^{-2}$(400 Å oxide)
[2] ICECREM N_2anneal, 900°C, 30 min
[3] experimental annealed profile

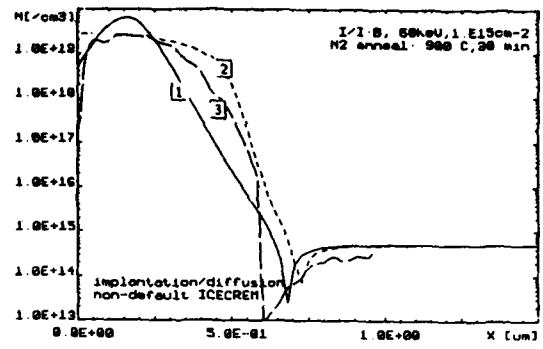


FIGURE 7 :

ICECREM non-default profiles vs experimental SRP profile :
[1] ICECREM implantation B, 60keV, $10^{14}cm^{-2}$(400 Å oxide)
[2] ICECREM N_2anneal, 900°C, 30 min
[3] experimental annealed profile

Simulations prove that due to the bevel angle used for SRP the measured junction is 20% less deep than the real junction in a non-beveled sample. This bevel-effect [7] is also responsable for the kink observed in the SRP-profile.

A more profound basis for the study and improvement of the diffusion model in the different simulators is established. Also this approach provides a more physically realistic strategy for adapting the default model parameters in order to fine tune the simulated results towards the experimental results.

It is made clear that it is very important to look carefully at the existing programs and the implemented models in order to get reliable results. Indeed, apparently good results can be obtained for certain process steps or subsequent process steps, but more critical conditions can cause a failure of the model.

ACKNOWLEDGEMENTS

The authors would like to thank the MAP/ARS group of IMEC for the SIMS and SRP data.

REFERENCES

[1] D.A.Antoniadis, R.W.Dutton, IEEE Journal of Solid State Circuits, vol SC14, No2, April 1979, p412

[2] C.P.Ho, J.D.Plummer, S.E.Hansen, R.W.Dutton, IEEE Transactions On Electron Devices, vol ED-30, No.11, November 1983, p1438

[3] H.Ryssel, K.Haberger, K.Hoffmann, G.Prinke, R.Dümcke, A.Sachs, IEEE Transactions On Electron Devices, vol ED-27, No.8, August 1980, p1484

[4] R.B.Fair, PREDICT, Microelectronics Center of North Carolina (MCNC)

[5] S.Selberherr, Analysis and Simulation of Semiconductor Devices (Springer-Verlag, Wien, 1984)

[6] W.Maes, K.M.De Meyer, L.Dupas, IEEE Transactions on CAD, vol.CAD-5, No.2, April 1986, pp.320

[7] J.Albers, Some Aspects of Spreading Resistance Analysis, in: Gupta,D.C. and Langer,P.H.(eds.), Emerging Semiconductor Technology, ASTM STP 960, (American Society for Testing and Materials, 1986)

SHALLOW JUNCTIONS OF BORON IMPLANTED IN GE^+ PREAMORPHIZED <100> SI WAFERS

A. La Ferla*, S. Cannavò°, G. Ferla°, V. Raineri*, E. Rimini*

* Dipartimento di Fisica, Corso Italia 57-I95129 Catania - Italy.
° SGS - Microelettronica S.p.A., Stradale Primosole 50, I95100 Catania Italy.

p-type shallow junctions in silicon were obtained by preamorphization with Ge^+ ions, 20 KeV B^+ implants at doses in the $5*10^{14}$ - $5*10^{15}/cm^2$ range, and thermal annealing at 850 °C or 950 °C for 1/2 hr. The junction depth was < 0.3 μm in the preamorphized wafers and > 0.3 μm in the B-bare Si implanted wafers. The leakage current density measured at a reverse bias of 10 v was about 100 pA/cm^2 for both procedures. The leakage maps on the 5" wafer gave a density of short circuits of $0.6/cm^2$.

1. INTRODUCTION

The formation of shallow junctions by ion implantation is hampered by channeling effect and by the partial electrical activation in the tail region of the implanted profile in particular for B-type region [1]. Preamorphization with different ions as Si^+, Ge^+, Sn^+ of a surface layer deeper than the active range of the p-n structure prevents channeling tails [2-3] and the use of rapid thermal annealing [4] provides a quite complete dopant activation.

The use of Ge^+ ions to preamorphize the layer presents some advantages [5]. The implantation can be performed at room temperature, Ge has an infinite solid solubility in Si, its large covalent radio causes a biaxial compressive strain which can compensate the lattice strain that arises after subsequent high dose implantation and annealing of impurities like B or P. The Ge implantation is superior to the Si^+ implantation in achieving uniform amorphization and a regrown region of high structural perfection results [6].

In the present work we report in detail the electrical characterization of pre-amorphized Si layers with Ge^+ ions, subsequently implanted with B^+ and then thermally annealed, in view of the relevance of the electrical response of shallow junctions for device applications.

2.EXPERIMENTAL

Silicon wafers, 5" n-type of 1.5 - 4.0 Ω *cm resistivity, of <100> orientation were preamorphized with Ge ions - 150 KeV to a fluence of $5*10^{14}/cm^2$. Some wafers were subsequently implanted with 20 KeV of B^+ to a fluence of $5*10^{14}/cm^2$ (R_p = 670 Å, DR_p = 340 Å). The annealing was performed in a furnace under N_2 flow at temperatures of 850 °C and of 950 °C for 1/2 hr.

The samples were analyzed by 2.0 MeV $^4He^+$ Rutherford backscattering in combination with channeling effect technique. The depth profile of carrier concentration and mobility was obtained by sheet resistence and Hall measurements in layer-by layer removal technique by anodic oxidation. The electrical behaviour was determined by the forward and reverse I-V characteristic and by the leakage maps.

3. EXPERIMENTAL RESULTS

The thermal regrowth of preamorphized Ge^+ layers and subsequently implanted with 20 KeV B^+ was investigated in the 500-600 °C temperature range. The initial amorphous layer amounts to 160 nm. The epitaxial regrowth rate of the amorphous layer is the same for Ge^+ and self-ion Si^+ implantation. The presence of B dopants at concentration of $5*10^{19}/cm^3$ enhances the rate of about a factor ten, in agreement with the data obtained [7] in self-ion implanted layers. Our data support the use of Ge instead of Si, in inducing an amorphous layer. After regrowth Ge atoms occupy substitutional lattice sites.

The sheet resistance of samples implanted with $5*10^{14}/cm^2$ B in either pre-amorphized or virgin <100> Si n-type substrates, and thermally annealed at different temperatures for 30' is shown in Fig. 1.

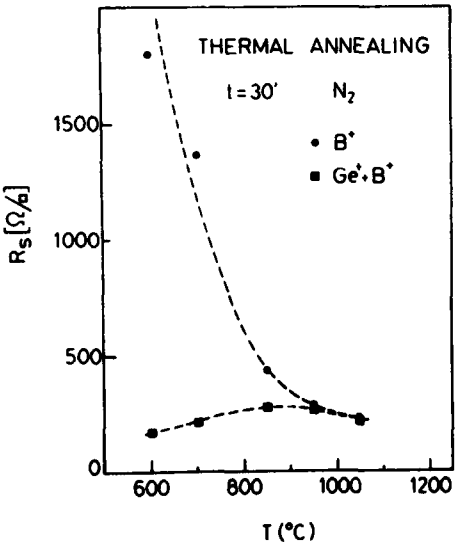


Fig. 1 - Sheet resistance versus annealing temperatures for $5*10^{14}/cm^2$ - 20 KeV B implanted into Ge preamorphized (■) and bare Si (●) respectively.

Sheet resistence is higher in B implanted samples than in B implanted

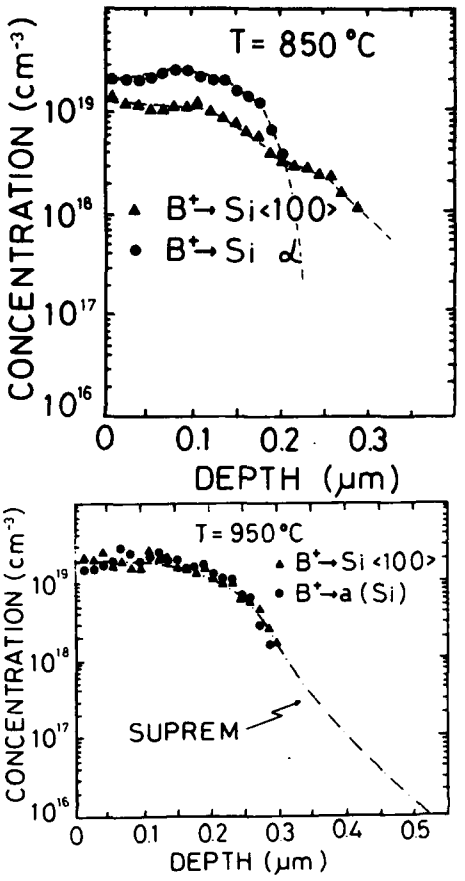


Fig. 2 - Carrier concentration profile of $5*10^{14}/cm^2$ - 20 KeV B^+ implanted into preamorphized (●) and bare (■) Si wafers after thermal annealing for 30' at 850 °C (a) and 950°C (b)

into preamorphized samples up to 900 °C annealing temperature. This behaviour is associated to the regrowth of α-layers and then to the electrical activation of the embedded B dopant at temperatures well below those required to anneal out point and extended defects present in the B implanted samples [8]. The sheet resistance of the layers amounts to 250 Ω/□ in agreement with SUPREM simulation.
The doping and mobility profiles were determined by etching off thin layers of Si one after the other and by measuring the sheet resistance and the Hall coefficient. The thin layers were

removed by anodic oxidation and the Van der Paw geometry was adopted for the electrical measurements. The carrier concentration profiles obtained by this procedure, are shown, in Fig. 2a and 2b for $5*10^{14}/cm^2$ - 20 KeV B^+ implanted into bare Si and into preamorphized <100> Si substrates after annealing at 850 °C and 950 °C respectively. After annealing at 850 °C the activated B^+ dopant fluence amounts to $4.4*10^{14}/cm^2$ in the preamorphized Si and to $6.9*10^{13}/cm^2$ in the bare Si samples in agreement with the resistivity data reported in Fig. 1. The mobility is of 60 $cm^2/v*s$ and 80 $cm^2/v*s$ respectively. After annealing at 950 °C the measured profiles nearly coincide in the investigated range. The activated dose amounts to $4.4*10^{14}/cm^2$ in the preamorphized and $4.3*10^{14}/cm^2$ in the bare Si respectively. The mobility is 60 $cm^2/v*s$ for both. It must be pointed out that the minimum carrier concentration detectable with accuracy by this method is of about $5*10^{17}/cm^2$, i.e. at least two orders of magnitude higher than the substrate doping. The location of the junctions is then determined by the shape of the concentration profiles at values lower than $10^{17}/cm^3$.

As an example the B^+ profile calculated by SUPREM is reported, as full line, in Fig.2b. The tail is approximated in the simulation by an exponential decay, whose slope depends on several parameters, as orientation of the wafer with respect to the beam, annealing temperature, etc. The first part of the simulated profile accounts quite well for the experimental data; but the distribution in the tail cannot be inferred by these measurements, and the location of the junction can be only guessed.

A staining technique has been adopted to measure the junction depth and the following results were obtained after 850 °C and 950 °C annealings: X_j(850°C) = 0.32 μm and 0.2 μm in the bare and in the preamorphized Si, X_j(950°C) = 0.35 - 0.4 μm and 0.28 - 0.3 μm in the bare and in the preamorphized Si. The accuracy of the method is about ± 0.05 μm. The 950 °C seems a good choice for the annealing temperature because at 850 °C a large amount of defects is still present and at higher temperature as 1000 °C for few seconds, a considerable diffusion of the dopant occurs.

The electrical behavior of the formed p-n junctions was investigated by measuring the forward and the reverse I-V characteristics and by the leakage maps in diodes of 0.22 cm^2 area. The reverse cheracteristics are shown in Fig. 3 for the adopted thermal procedures. No pregettering was adopted for the processed diodes. The annealing at 950 °C results in a quite reasonable electrical behavior. The leakage current density measured at a reverse voltage of 10 v was about 100 pA/cm^2 for both procedures (implantation in bare Si or in preamorphized Si). This low value is very close to the theoretical one and indicates that defects and generation centres in the depleted p-n junction are practically absent. The forward I-V characteristcs, not shown, are fitted by an ideality factor $m \backsim 1.02$ in the exponential factor.

The leakage current maps were also measured in the processed 5" wafer at a reverse bias of 15 v. A typical map is reported in Fig. 4, together with the histogram of the distribution. The density of short circuits amounts to $0.6/cm^2$. The wafer was processed in such a way that a central region of 2" diameter was preamorphized with Ge and then all the wafer was implanted with 20 KeV B^+. A direct comparison is then possible between preamorphized and as implanted diodes.

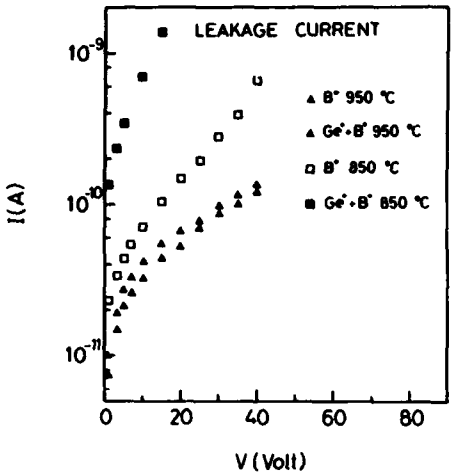


Fig. 3 - Reverse I-V characteristics of 0.22 cm^2 diodes implanted with 20 KeV B^+ - $5*10^{14}/cm^2$.

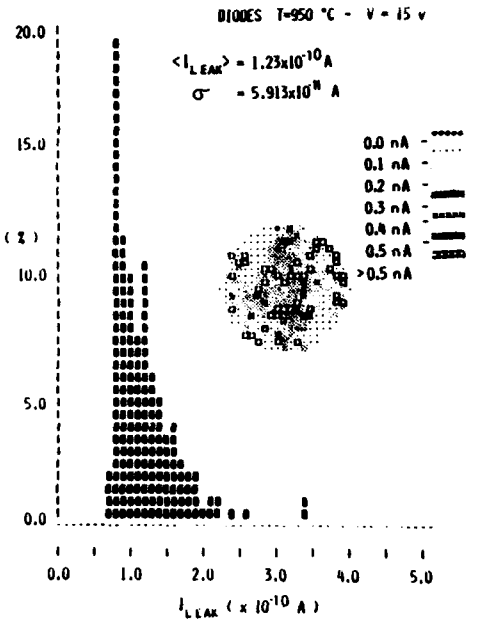


Fig. 4 - Histogram and map of leakage current at - 15 v of diodes annealed at 950 °C - 30'; the two inches diameter central region was preamorphized with Ge^+ implantation.

The leakage map and the distribution indicate that within the experimental accuracy no significant difference exists in the electrical behavior of the diodes processed in the two ways. This indicates that the preamorphization with Ge doesn't deteriorate with the residual defects the electrical response.

ACKNOWLEDGEMENTS

The authors wish to tank Dr. M. L. Polignano and Dr. G. F. Cerofolini of the SGS-Microelettronica S.p.A. - Agrate (MI), for an helpful discussion of the electrical results.

Work supported in part by " Programma Nazionale per la Microelettronica ".

REFERENCES

[1] C. Hill, Nucl. Instr. Meth. Phys. Res. B19/20, 345 (1987).

[2] C. Carter, W. Maszara, D. K. Sadana, G. A. Rozgonyi, J. Liu and J. Wortman, Appl. Phys. Lett. 44, 459 (1984).

[3] A. C. Ajmera and G. A. Rozgonyi, Appl. Phys. Lett. 49, 1269 (1986).

[4] G. K. Caller and T. E. Seidel, Appl. Solid State Science-Suppl. 2c editor D. Kahng, Academic Press N. Y. 1985 p. 2.

[5] D. K. Sadana, E. Myers, J. Liu, T. Finstad and G. A. Rozgonyi, Mat. Res. Soc. Symp. Proc. 23, 303 (1984).

[6] S. Prussin, E. R. Weber, K. S. Jones, Nucl. Instr. Meth. Phys. Res. B21, 496-502 (1987).

[7] J. S. Williams in "Surface Modification and Alloying" edited by J. M. Poate, G. Foti and D. C. Jacobson, Plenum Press, N. Y. 133 (1983).

[8] W. K. Hofker, Philips Research Reports, Suppl. N. 8 (1975).

THE EFFECT OF HIGH PRESSURE STEAM OXIDATION ON PHOSPHORUS DIFFUSION IN SILLICON

WU BAI LU

Graduate School of Academia Sinica, China

ZHANG AI ZHEN

Beijing Institute of Semiconductor Devices, China

LI SHY LIN

National Institute of Metrology, China

XUE SHI YING

Academia Sinica, China

The effect of high pressure steam oxidation (7.5-10.7 atm.) on phosphorus extrinsic diffusion in (111)--and (100)--silicon at 700°-970°C has been examined by spreading resistance probes and ellipsometery. It has been found that the OED and ORD appear at the higher and the lower temperature, respectively. The OED-ORD transition point is at about 880°C for 40' in 7.5 atm. for (111)-silicon. The difference in effective diffusion coefficients between oxidation and non-oxidation regions (D) is proportional to $(x_0/t)^n$, the power figure n is related to the oxidation conditions. These results can be explained satisfactorily by considerations which take into account oxidation rate and concentration effect on phosphorus diffusion in silicon.

With the development of VLSI, the smaller device dimensions are needed. Therefore the fine control of impurity profiles is more important. The high pressure oxidation is an excellent oxidation process for VLSI. A precise understanding of the effect of this oxidation on impurity redistribution is a requirement for device processes and has important significance for further solving the physical mechanism of oxidation enhanced diffusion (OED) and oxidation retarted diffusion (ORD).

This paper reports that several new phenomena found under the high pressure oxidation are different from those in atmospherical oxidation.

1. EXPERIMENTAL CONDITIONS

Original silicon wafers were 7-15Ω·cm, p type (100)- and (111)- orientated single crystal silicon. The phosphorus diffusion were performed with sp_1 planar solid source at 1140°C for 8'. After that the surface phase layers were removed with the dilute HF solution. Then the Si_3N_4 films of 1000 Å thickness were deposited on the silicon surface and subsequently were photolithographically patterned into 1 m.m. wide parallel stripes so that both the nonoxidizing and oxidizing regions could be on the same wafer. The thermal oxidation was carried out in a high pressure steam oxidation system which was automatically controlled by a microprocessor. Before and after oxidation, the spreading resistance profiles (R-X) were measured by ASR-100B spreading resistance probes. The oxide thickness (x_0) were measured by a TP-77 ellipsometer.

2. EXPERIMENTAL RESULTS AND DISCUSSIONS

The oxidation curves of (111)-Si in 7.5

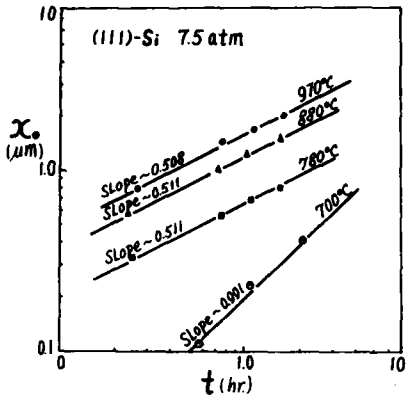


FIGURE 1
Oxidation curves of silicon in high pressure steam at 700^0–970^0C

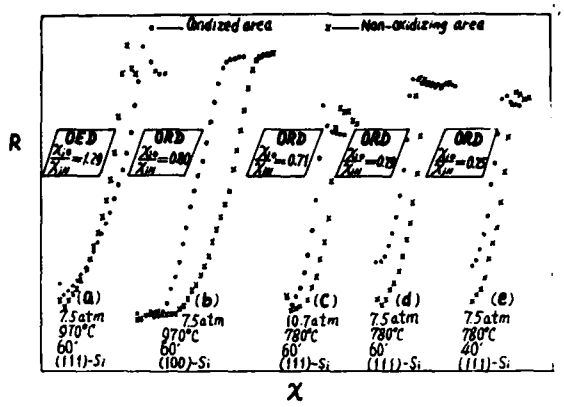


FIGURE 2
Matched spreading resistance profiles of phosphorus in silicon

atm. steam at 700^0–970^0C are given in Fig.1. It can be seen that the oxidation rate (dx_0/dt) is very fast at 970^0, 880^0, 780^0C, it has run in the parabolic oxidation regime. While at 700^0C, it is in the linear oxidation regime. It is implied that the difference in impurity diffusion in silicon under high pressure and atmospherical oxidation is not due to the difference in the oxidation equation, but is correlative with the oxidant supply.

Several typical spreading resistance profiles for different experimental conditions are shown in Fig.2. From 2a and 2b, it can be observed that the impurity distribution and the ratio of the junction deepth, x_{jo}/x_{jn}, are related to the crystalline orientation of samples. Within the pressure and the temperature range used in this experiment, all of the phosphorus diffusion in (100)-Si are retarded (ORD), and its diffusion coeffecients are all bigger than those of (111)-Si at the same condition. By comparing 2c with 2d, it is obvious that the effect of oxidation on diffusion is correlative with pressure (p) of oxidant.

A plot of the logarithm of diffusion coefficient versus the reciprocal of absolute temperature is shown in Fig.3 for phosphorus in (111)-silicon in 7.5 atm. Here, the diffusion coefficients D_0, D_N were calculated from the measured junction depths x_j and sheet resistances in both the oxidation and the nonoxidation region according to Gaussian distribution. By doing so, the segragation effect of the moving boundary in the oxidation region is ignored. The validity has been discussed by some authors [1, 2, 3]. It can be noted from this figure that $D_0 > D_N$ at the higher temperature (970^0C), that is OED. While $D_0=D_N$ at 880^0C for 40', it is the transition point of OED-ORD. When the oxidation time t is over 40' or the temperature is lower than 880^0C, then $D_0<D_N$, that is ORD. Another regular phenomenon is that D_0 decreases as increasing the oxidation time t at the same temperature, and D_N also decreases except for 970^0C so that the enhancement decreases while the retard increases.

A plot of ΔD versus time-mean effective oxidation rate x_0/t is shown in Fig.4. A result of normal dry oxygen oxidation given by Y. Ishikawa [9] is also plotted in the figure. From Fig.4, ΔD and x_0/t have a power relationship except for near the transition point.

$$|\Delta D| = A(x_0/t)^n$$

within the parabolic oxidation region indicated in Fig.1.

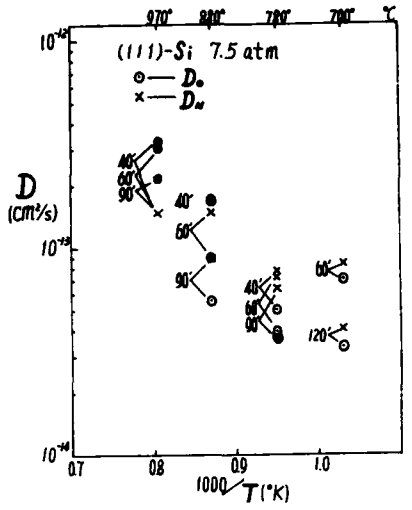


FIGURE 3
Diffusion coefficient of phosphorus in silicon versus reciprocal temperature

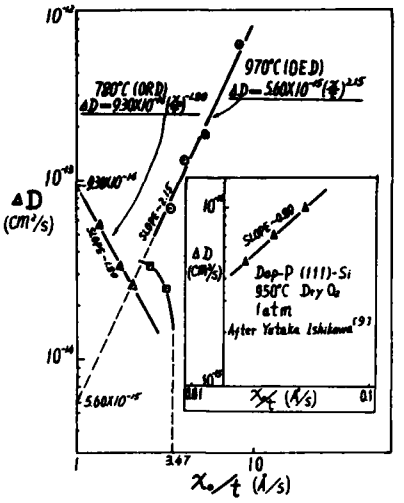


FIGURE 4
Oxide growth rate dependence of D of phosphorus in silicon

$$x_0/t = 2(dx_0/dt)$$

Thus, we have

$$|\Delta D| \propto (dx_0/dt)^n$$

Where, the n is different from the predictive value (n=1) of S.M.Hu model [7] and the results given at normal oxidation by other authors [9, 10]. The n can be positive or negative. It is depend on the temperature, pressure, oxidant composition and crystalline orientation of samples. Particularly near the OED-ORD transition point, the n approaches infinity as a limit. It means some criticality and is well worth notice.

Another interesting phenomenon is the relationship of $D_0/D_N - \frac{1}{T}$ plotted in Fig.5. The right half is our result in 7.5 atm., but the left half was obtained in dry O_2 by Francis and Dobson [11]. When the temperature is increasing from 700°C to near 1300°C, it undergoes two transitions from ORD to OED and again from OED to ORD. By comparing it with the relationship of the length of oxidation induced stacking faults (OSF) Vs 1/T of Fig.6, it can be found immediately that OSF also undergoes two transitions from the shrinkage to the growth, again to the shrinkage in the similar temperature range. These results verify once again that OED (ORD) and OSF are two behaviours coming from a same physical process - the injection of non-equilibrium point defect, and the phosphorus diffusion in silicon is mainly via interstitialcy mechanism.

The measured effective diffusion coefficient (D) may be considered to be a superposition of three physical effects: the Fick effect Di, the concentration effect ΔDc and the surface oxidation effect ΔD_0, that is

$$D = Di + \Delta Dc + \Delta D_0$$

In the nonoxidation region, ΔD_0=0. Also $\Delta Dc(970^0C)$ =0, since the surface impurity concentration ($1.3\times10^{19}cm^{-3}$) is lower slightly than the intrinsic carrier concentration in silicon at 970°C. Thus, $D_N(970^0C)$=Di and has no change with t. While below 970°C, ΔDc≠0 and change with t, and so do the D_N. In the oxidation region, there are both ΔDc and ΔD_0. Strictly speaking, ΔDc and ΔD_0 effect each other [4]. But, as a first approximation, we can assume [5, 6] $\Delta D_c(0) \simeq \Delta D_c(N)$, then

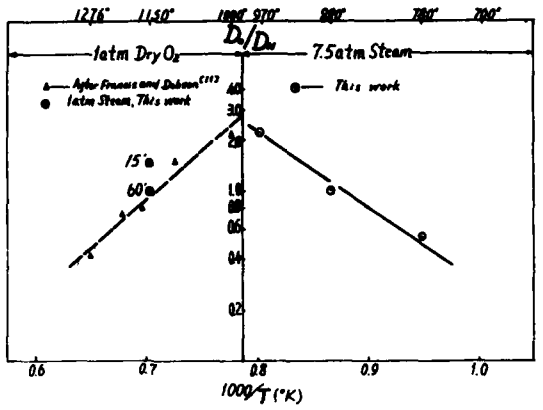


FIGURE 5
Ratio of the diffusion coefficients (D_O/D_N) versus reviprocal temperature

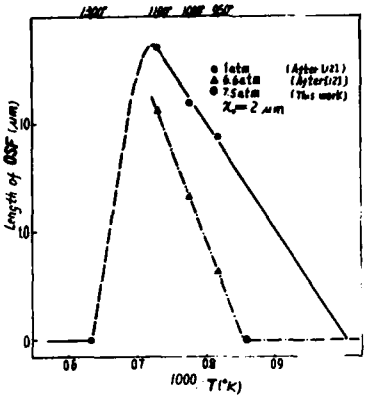


FIGURE 6
Length of OSF versus reciprocal temperature

$\Delta D=D_O-D_N=\Delta D_O$, that is, the measured ΔD can be approximately considered to be the result from oxidation solely.

The ΔD_O, as has been recognized [7, 8], come from non-equilibrium point defect injection into silicon caused by oxidation reaction at SiO_2-Si interface. In turn, the point defect injection would be dependent upon the relative rate between the diffusion transporting of oxidant at SiO_2-Si interface and the reaction consuming of it at the interface. This suggests that the lack and the excess of oxidant at the interface result in interstitail and vacancy injection, and so for OED and ORD of phosphorus. The pressurizing is in favour of oxidant transporting, the decrease of temperature and the increase of time result in retard of interface reaction, hence ORD is favoured. In a similar way, the difference in oxidant composition and crystalline orientation also has effects on OED (ORD).

In summary, the result of high pressure oxidation shows that the phosporus diffusion can be enhanced or retarded, it is dependent upon the oxidation condition.

REFERENCES

[1] Wu Bai-Lu et.al., ICSICT'86, Beijing, p.70

[2] T.Kato et.al., Jpn, J.Appl. Phys. 3(7), 377 (1964).

[3] Y.Nabeta et.al., J.Electrochem. Soc.,123, 1416 (1976).

[4] B.E.Deal et.al., J.Electrochem. Soc.,112, (4), 430 (1965).

[5] C.P.Ho et.al., J.Electrochem. Soc., 126(9), 1516 (1979).

[6] Wu Bai-Lu et.al., to be published.

[7] S.M.Hu, J.Appl. Phys. 45(4), 1567 (1974).

[8] T.Y.Tan et.al., Appl. Phys. Lett. 40(7), 1. April, 1982.

[9] Yutaka Ishikawa et.al., J.Electrochem. Soc. 129(3), 644, 1982.

[10] K.Taniguchi et.al., J.Electrochem. Soc., 127(10), 2243, 1980.

[11] R.Francis and P.S.Dobson, J.Appl. Phys. 50(1), 280, 1979.

[12] Natsuro Tsubouchi et.al., Jpn.J.Appl. Phys. 17, 223-228 (1978).

SHALLOW JUNCTION FORMATION USING $CoSi_2$ AS A DIFFUSION SOURCE

V. Probst[1], P. Lippens, L. Van den hove, K. Maex, H. Schaber* and R. De Keersmaecker

Interuniversity Microelectronics Center (IMEC v.z.w.), Kapeldreef 75, B-3030 Leuven, Belgium
*SIEMENS AG, Central Research and Development, Otto-Hahn-Ring 6, D-8000 Munchen 83, BRD

Thin layers of $CoSi_2$ (120 nm) were used as a source for B and As diffusion in order to form shallow steep junctions with high interface concentration. SIMS depth profiling as well as two-dimensional characterisation of the indiffusion demonstrate the power fo this technique over a wide range of temperatures and times. Diodes with a high yield and a very low leakage current density ($\simeq$1 nA/cm^2) prove the reliability of the process.

1. INTRODUCTION

The shrinking of device dimensions in integrated circuits necessitates a significant reduction of junction depths towards the 0.1μm scale. In order to improve device and circuit speed, however, parasitic elements such as series resistance have to be reduced as well. These two requirements cannot be satisfied simultaneously by conventional junction formation processes, nor by dopant diffusion from poly-Si such as applied e.g. in self-aligned bipolar devices [1]. Silicides, on the other hand, are proven to be successful in order to reduce the contact and sheet resistances. The silicidation of extremely shallow diodes, however, requires special attention. When the conventional silicidation of preformed shallow junctions is applied, the diode yield decreases drastically with decreasing junction depth/silicide thickness ratio [2]. Therefore, the diffusion of dopants from the silicide into the silicon is an attractive alternative to this process [3]. Besides the benefit of avoiding direct implantation of dopants into the single-crystal silicon, the proposed method of junction formation is expected to be 'self-adjusting' to the silicide/silicon interface shape which has a certain degree of roughness. Therefore, junction shortage due to inhomogeneous silicidation can be avoided.

2. EXPERIMENTAL

In this work, $CoSi_2$ was used as a source for B and As diffusion into mono-Si. In order to study the fundamental characteristics of this diffusion source, the $CoSi_2$-salicide process, as described in [2], was applied first on unpatterned Si-wafers. After formation of 120 nm $CoSi_2$ at 700°C for 30 s (RTP), the desired dopant was implanted into the silicide with a dose of 5.E15 cm^{-2}. Simulations using the program TRIM85 [4] were used to determine a suitable implantation energy assuring confinement of the implant within the $CoSi_2$-layer (20 keV in the case of B, 50 keV for As). Prior to the diffusion step, the silicide was capped with 200 nm of CVD-SiO_2 in order to prevent dopant loss to the ambient. The diffusion cycle was either performed in a conventional furnace or in a rapid thermal processing system (RTP), in order to test the diffusion behaviour and stability of the $CoSi_2$-source over a wide range of temperatures and times.

In a second experiment shallow silicided n^+-p and p^+-n diodes were fabricated using the diffusion of dopants from $CoSi_2$. Active areas were defined by a conventional LOCOS-technique. The silicidation, implantation and diffusion steps were carried out following identical conditions as for the unpatterned wafers. Then a CVD-oxide was deposited and contact windows were opened. Ti/W was used as a diffusion barrier between the silicide and the Al-metal layer. Finally, a sintering in forming gas at 450°C for 30 min was carried out.

3. RESULTS

Four-point probe measurements were performed between the processing steps to check the change in $CoSi_2$ sheet resistance. Due to implantation damage, the initial sheet resistance

[1]Permanent address: Siemens AG, Central Research and Development (München)

of 1.25 Ω/□ after silicidation [2] increased to 5 Ω/□ in the case of B implantation and 3 Ω/□ for As. The subsequent diffusion steps anneal this implantation damage and cause the complete recovery of the $CoSi_2$ sheet resistance (e.g. 1.1 Ω/□ after 950°C anneal for 30 min), which is in good agreement with [5].

The diffusion of the dopants in the silicide was measured by secondary ion mass spectrometry (SIMS). Figure 1 shows the as-implanted B profile in comparison with the 800°C/30 min diffusion step. Due to the fast diffusion of B in $CoSi_2$ at that temperature a complete equidistribution has taken place in the silicide at a concentration level of about 3.E20 cm^{-3}. At the interface to the SiO_2-layer, boron tends to segregate (SiO_2 was etched off prior to SIMS measurement).

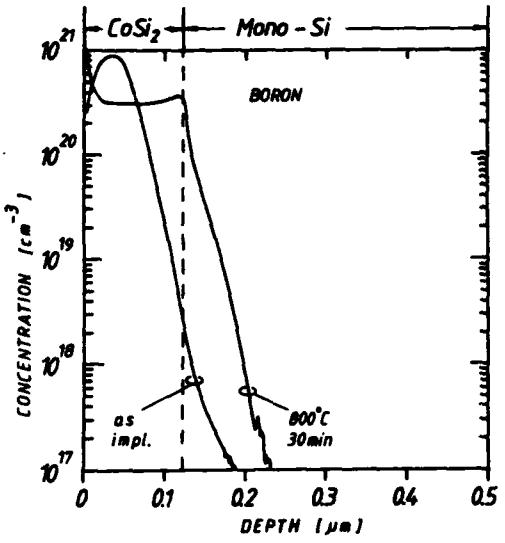


FIG. 1: Boron profiles measured by SIMS: as-implanted and diffused out of $CoSi_2$ at 800°C for 30 min.

In fact, more important for device applications is the indiffusion of the dopants from the silicide into the mono-Si. Especially the temperature-dependent junction depth and interface-concentration are of main interest for scaling and contact-resistance respectively. In order to exclude matrix-effects during SIMS-measurement and to avoid limitations in depth resolution, the $CoSi_2$ was selectively removed from the mono-Si by etching in 25 % HF. Figure 2 shows the indiffusion of boron from $CoSi_2$ for different heat-cycles. Junction depths from 100 nm (for 800°C/120 min) to 500 nm (for 1100°C/60 s) and

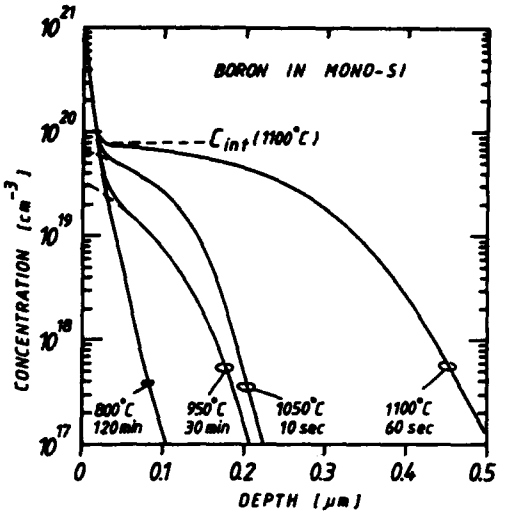


FIG. 2: SIMS-profiles of B-indiffusion from $CoSi_2$ at different heat cycles.

interface-concentrations between 3.E19 cm^{-3} (at 950°C) and 8.E19 cm^{-3} (at 1100°C) show the reliability of $CoSi_2$ as a diffusion source over a wide temperature and time range. A junction depth of about 200 nm can either be achieved by 950°C/30 min furnace anneal or by 1050°C/10 s RTP. As expected, the surface concentration is higher for the RTP sample. The same tendency is shown in fig. 3 for As-indiffusion. However, due

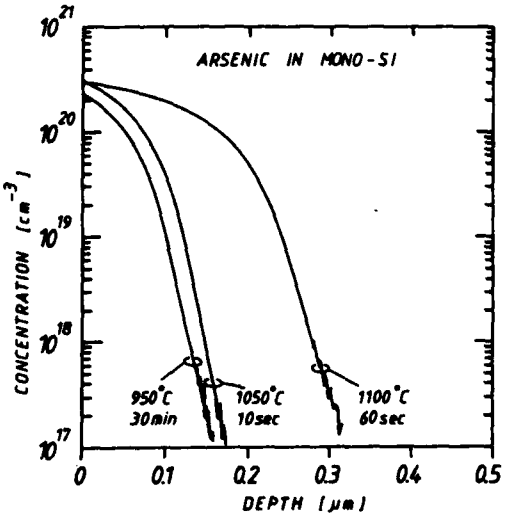


FIG. 3: SIMS-profiles of As-indiffusion from $CoSi_2$ at different heat cycles.

to the lower diffusivity of As in mono-Si [7], the As-junctions are shallower than the B-junctions formed at equivalent conditions. Interface concentrations between 2.E20 cm^{-3} and 3.E20 cm^{-3} were achieved for As at 950°C and 1100°C, respectively. In general, the indiffusion behaviour of B and As from $CoSi_2$ is very similar to the indiffusion from poly-Si [6], [7]. Slight deviations in the interface concentrations could be due to the different segregation behaviour or solid solubility of dopants in $CoSi_2$ which requires additional experiments.

SIMS-depth profiling on shallow junctions (< 50 nm) (e.g. 800°C/120 min As-indiffusion) need special measurement conditions such as low energy primary ion beam, which increases measuring time drastically. For determination of the junction depth only, the method of bevel and staining (BS) is more efficient on these samples. Table 1 gives a summary of the junction depths obtained for B and As for various annealing conditions, measured with SIMS, BS and spreading resistance probe (SRP). Taking into account that these independent methods yield junction depths at different doping levels, a good agreement is obtained between the different results which proves the reliability of the data.

	ARSENIC					BORON				
DIFFUSION STEP	800°C 2 h	900°C 30'	950°C 30'	1050°C 10"	1100°C 60"	800°C 2 h	900°C 30'	950°C 30'	1050°C 10"	1100°C 60"
X_{jSIMS} (nm) at 1.E17 cm^{-3}			150	170	310	110	150	200	220	500
X_{jBS} (nm)	40	70	140	120	290	110	130	190	190	430
X_{jSRP} (nm)			160		330			110		470

TABLE 1: Comparison of junction depths obtained by indiffusion from $CoSi_2$ and measured with SIMS, SRP, bevel and staining (BS).

An important issue is the lateral homogeneity of the junction. Since the mentioned analysis methods give one-dimensional information (averaged over a large sample area) only, SEM was applied on selectively etched cross-sections to visualize the $CoSi_2$-grains and the two-dimensional shape of the diffusion front with high resolution (at a dopant concentration of about 5.E18 cm^{-3}).

Figures 4a and 4b give examples for the case of B-diffusion at 900°C/30 min and 1050°C/30 s respectively. The $CoSi_2$-

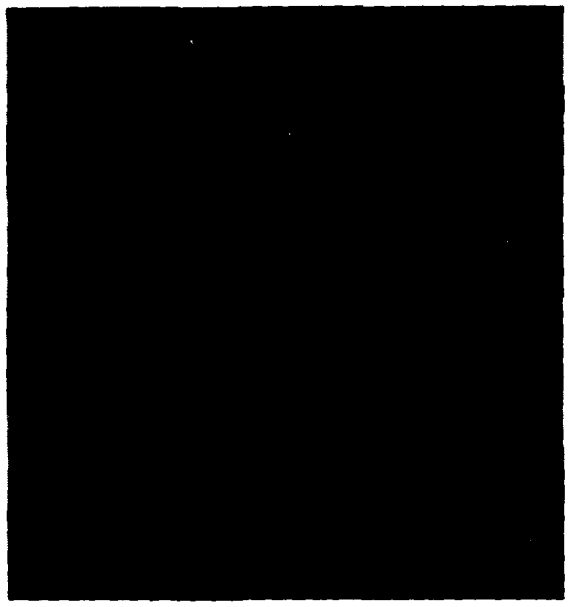


FIG. 4a: SEM-cross-section showing the $CoSi_2$/Si interface and the diffusion front for the boron diffusion at 900°C, 30 min. FIG. 4b: SEM-cross-section for boron indiffusion at 1050°C for 30 s. FIG. 4c: SEM-cross-section showing $CoSi_2$-'globules' surrounded by the diffusion front for As-diffusion at 1100°C, 60 s.

grains, the interface to the mono-Si and the diffusion region are clearly seen. These micrographs reveal that the diffusion front follows the $CoSi_2$/Si interface at a nearly constant distance. The explanation for this self-adjusting mechanism is similar to that for the poly-Si diffusion source in ref. 7. The $CoSi_2$/Si interface itself behaves like a grain boundary with high diffusivity. The dopants are mainly supplied by the vertical grain boundaries, but also from the bulk of the $CoSi_2$ [8]. Reaching the interface, they immediately redistribute in the lateral grain boundaries before slowly diffusing into the mono-Si.

Figures 4a-4b also show that the lateral growth of the $CoSi_2$-grains from $d_G \simeq 0.3$ μm at 900°C/30 min to $d_G \simeq 0.6$ μm at 1050°C/30 s is accompanied by a local conglomeration of the silicide. This in consequence causes a temperature-dependent local change in thickness and an increase of the interface-roughness from about 50 nm at 900°C/30 min to about 140 nm at 1100°C/60 s. This is even more pronounced in the case of As-doping, where the $CoSi_2$-layer has balled up locally after a heat cycle of 1100°C/60 s. However, even this extreme case confirms the self-adjusting mechanism of indiffu-

sion by showing the $CoSi_2$-'globules' surrounded by the diffusion front.

Figure 5 gives a typical I-V (forward and reverse) characteristic of diodes formed by a diffusion of B or As at 800°C/120 min in N_2. Table 2 summarizes the average leakage current density for several diffusion conditions (devices with leakage current density higher than 20 nA/cm^2 were considered defective). Even for the highest temperature cycles (1100°C) where the silicide has balled up locally (fig. 4c) leakage currents as low as 1 nA/cm^2 (at 5 V reverse bias) were observed. This also indicates that the generation of trap centers related to Co in the space charge region must be much less pronounced than expected from diffusion length (about 160 μm) and solid solubility (about 1.E15 cm^{-3}) [9] of Co in Si at 1100°C/ 60 s. Due to the non-planar shape of the junction (fig.5) , the breakdown voltage is expected to decrease [10]. However, no difference was seen in comparison with the values obtained from the non-silicided control devices (20 V for As, 30 V for B).

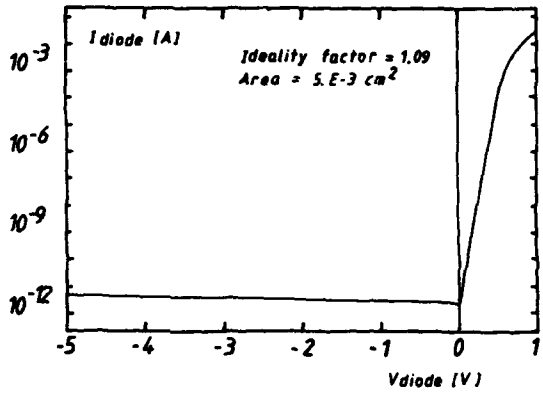


FIG. 5: I-V characteristic of n^+/p-diode: As diffused out of $CoSi_2$ at 800°C for 120 min.

	ARSENIC		BORON	
	mean [nA/cm^2]	% good devices	mean [nA/cm^2]	% good devices
800°C, 120 min	0.9	100	3.3	100
850°C, 60min	1.3	96	3.6	100
950°C, 30min	0.9	100	2.4	96
1100°C, 10 s	3.8	80	1.8	100

TABLE 2: Comparison of leakage current and yield for B and As-diodes formed by indiffusion from $CoSi_2$.

4. CONCLUSIONS

It is demonstrated that $CoSi_2$ is a very useful and flexible diffusion source for arsenic and boron. SIMS-analyses show that shallow as well as deep junctions with high interface concentrations can be obtained by furnace anneal and RTA. Two-dimensional analyses of the diffusion (SEM) reveal an interface-related self-adjustment of the diffusion front which avoids junction shortage due to inhomogeneous silicidation. Diodes formed by B or As indiffusion from $CoSi_2$ show ideal forward and reverse characteristics with very good yield even for extremely shallow junctions of 40 nm (As) and 100 nm (B).

ACKNOWLEDGEMENTS

The authors would like to thank P. Eichinger and W. Vandervorst for SIMS-analyses, K. Wittmaack for useful discussions and also R. De Koninck for device-measurements. P. Lippens is indebted to the Belgian Institute for Scientific Research in Industry and Agriculture (IWONL), whereas L. Van den hove and K. Maex are supported by the Belgian National Fund for Scientific Research (NFWO).

REFERENCES.

1. A. W. Wieder, IEDM Tech. Dig. 1986, p.8
2. L. Van den hove et al., IEEE Trans. El. Dev., ED-34, (March 1987) 554
3. R. Liu et al., presented at the ECS-meeting, May 10-15 1987 (to be published)
4. J. P. Biersack and L. G. Haggmark, Nuclear Instruments and Methods 174, (1980) 257
5. B. M. Ditchek et al., presented at the MRS-meeting, april 1987 (to be published)
6. H. Schaber et al., J. Appl. Phys. 58, (1985) 11
7. V. Probst et al., Semiconductor silicon, (1986) 594
8. P. Gas et al., to be published in J. Appl. Phys.
9. H. Kitagawa et al., Jap. J. Appl. Phys. 21, (1982) 276
10. A. S. Grove, Physics and Technology of Semiconductor Devices, Ed. J. Wiley New York, (1967) 195

2-D EFFECTS DURING ISOLATION PROCESS: EXPERIMENTS AND SIMULATION

A. Seidl

Institut für Festkörpertechnologie
Paul Gerhardt Allee 42
D-8000 München 60

V. Huber

Siemens AG, ZT ZFE FKE 41
Otto-Hahn-Ring 6
D-8000 München 83

SEMIROX bird's beaks were processed for different sets of parameters including variation of the buffer oxide thickness, nitride thickness and temperature. A numerical simulator is used to discuss two-dimensional stress-and diffusion effects.

1. INTRODUCTION

For many standard oxidation processes it can be observed that the oxidation rate is strongly stress-dependent. A qualitative description of this effect has been given in [1,2]. The first effort for a quantitative analysis of this effect was made in [4] where circular etched silicon structures were oxidized and the oxide thickness was measured as a function of the radius. By using rotational symmetry the actual two-dimensional problem was reduced to one dimension and thus could be described by an ordinary differential equation. In [4] the different oxidation rates for concave and convex corners were attributed mainly to a pressure dependence of the viscosity thus allowing no conclusion on the behaviour of oxidant diffusion under stress.

As will be shown in the next section the length-to-width ratio of a bird's beak is a function of the diffusion ccefficient/reaction constant ratio. Thus the bird's beak experiment provides better orthogonality between the oxygen diffusion and reaction because it delivers as results not only the oxide thickness but also the length-to-width ratio of the under-diffusion region.

2. 2D DIFFUSION EFFECTS

In [3] it was shown by a simple qualitative model that the length of the bird's beak grows with increasing buffer oxide thickness and diffusion constant and drops with increasing reaction rate. In this case the under-diffusion of the mask via the buffer oxide was considered.

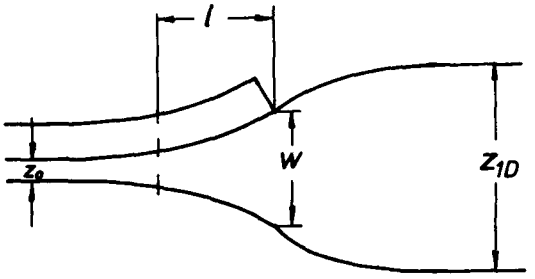


FIG. 1: Typical bird's beak geometry. The start of the under-diffusion-region is defined to be the position where the oxide thickness reaches $z_0+z_{1D}/10$. The width of this region is measured at the mask edge.

Exact numerical values for the dependence of the bird's beak length on D, K, and z_0 can be extracted from numerical simulation only. For this purpose a geometrical definition of the width-to-length (w/l) ratio was defined according to Fig. 1. For the numerical simulation the Deal-Grove model was ex-

tended to two dimensions by using the Finite Element method. Viscous flow was assumed for oxide deformation.

Fig. 4 shows that the dependence of the w/l-ratio on buffer oxide thickness is linear whereas it's dependence on the reaction/diffusion (k/D) ratio exhibits a rather nonlinear behaviour (Fig. 2).

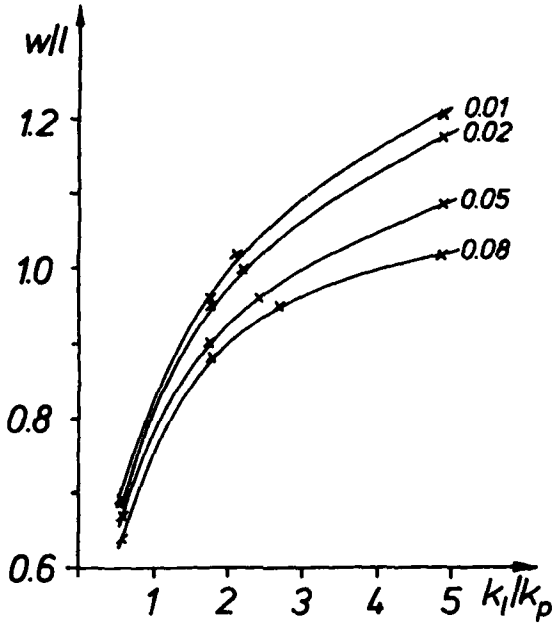


FIG. 2: The w/l ratio of a bird's beak depending on the reaction rate/diffusion coefficient ratio. (The linear and parabolic growth rates are related to these constants by $k_l/k_p = k/2D$)

3. STRESS EFFECTS

The shape of the bird's beak is influenced by mechanical stress. The two mechanisms involved are:
- direct deformation of the oxide by pressure exerted by the nitride mask
- stress-dependent coefficients (Diffusion, reaction, viscosity etc.)

The second set of mechanisms has been observed by various workers but a quantitative model has been given only by Kao [4]. The coefficients of the governing equations are dependent on mechanical stress by a Boltzmann type relationship.

(1) $$k_s = k_o \exp(-\sigma_n V_k/kT$$
$$D = D_o \exp(-pV_D/kT)$$
$$C^* = C^*_o \exp(-pV_c/kT)$$
$$\mu = \mu_o \exp(\alpha(T)p)$$

σ_n denotes the normal stress at the $Si-SiO_2$ interface, p the hydrostatic pressure, C^* the saturation concentration and μ the viscosity of the oxide. The V's are denoted as activation volumes.

4. EXPERIMENTS

In this work <100> lightly n-doped (5-20 Ωcm) silicon wafers were cleaned and oxidized in dry O_2 to form a buffer oxide between 100 and 800 Å. Silicon nitride of a thickness between 200 and 2000 Å was deposited. After chemically etching the nitride a wet oxidation was performed at temperatures between 900 and 1100 °C. The TEM preparation of the cross section was performed by mechanical lapping and successive ion etching.

5. VARIATION OF PAD OXIDE THICKNESS

Samples were fabricated with various pad oxide thicknesses at 1000°C oxidation temperature with a 1200 Å nitride mask. Qualitatively the profiles behave as expected and show reduced under-diffusion for lower pad-oxide thicknesses. However a quantitative comparison between measurement and stress-free simulation shows that the measured bird's beaks have the tendency to be longer



a



b

FIG. 3: TEM micrographs of bird's beaks grown with 800 Å (a) and 100 Å (b) buffer oxide together with stress-free simulation.

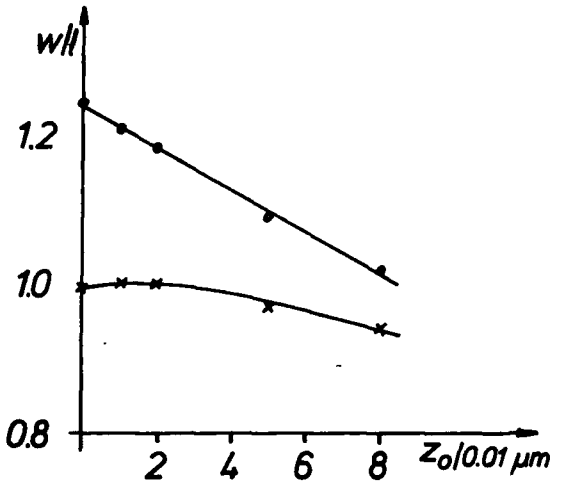


FIG. 4: w/l ratio versus buffer oxide thickness: x=measured, o=stress-free simulation

and flatter than the simulated profiles. The agreement is quite good for a large pad oxide thickness (Fig. 3a). However for small pad oxide thicknesses the difference is clearly visible (Fig. 3b). In Fig. 4 the w/l ratios of measured and simulated bird's beaks are plotted against the buffer oxide thickness. The comparison with the stress-free simulation case shows that the differences increase with shrinking buffer oxide thickness.

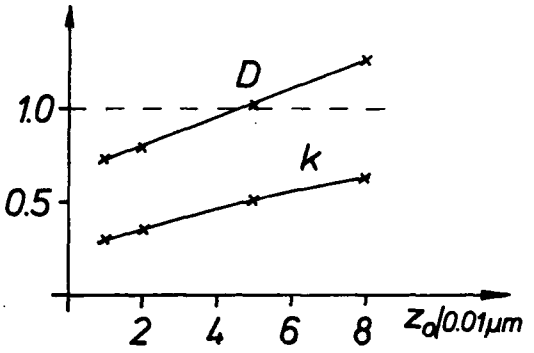


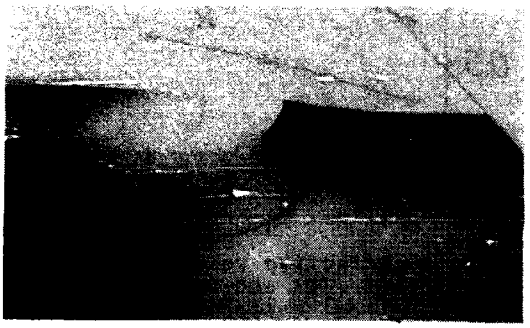
FIG. 5: Effective values for diffusion coefficient (D) and reaction rate (k) depending on buffer-oxide thickness normalized with respect to the stress-free values.

To quantitatively estimate the influence of stress on the diffusion- and reaction rates the bird's beaks were re-calculated with the values for these rates adapted such that the best fit for the geomtry of the under-diffusion-region was obtained. The values, which can be interpreted as mean-values for the whole under-diffusion-region, are shown by Fig. 5. The mask effect appears to be strong for a small pad oxide thickness. In this case the mask is strongly bent. If the values of the coefficients are translated into stress-values via Kao's model (1) it can be concluded that for thin pad oxides the mask exerts strong normal stress along the interface and that

compressive hydrostatic pressure is dominating throughout the under-diffusion-region. The mask effect becomes weaker for thick pad oxides. The mask is only slightly bent and seems to enhance the region of tensile pressure shown by numerical calculation. Nonetheless a reduced value of the reaction rate can still be observed.

7. Variation of Nitride Thickness

To investigate mechanical effects samples were fabricated with different nitride thicknesses at different temperatures. A direct deformation of the oxide, which leads to an assymetric profile could only observed for the maximum nitride thickness (2000Å) at 1100° (Fig. 6a). For the same nitride thickness, but lower temperature this effect is much less pronounced (Fig. 6b). This is due to the fact that the oxide becomes softer as compared to the nitride for high temperatures. In the second case the mask acts only indirectly through stress-induced reduction of the reaction rate. This results in stronger under-diffusion and thus leads to a very long bird's beak.



a



b

FIG. 6: Effect of nitride mask, 2000 Å: direct deformation at 1100 °C (a), influence via reduced reaction rate at 950° (b)

CONCLUSION

Extensive experimental data concerning the influence of technology parameters on the shape of MOS field oxide were obtained and characterized. Oxide growth is strongly influenced by mechanical effects. It was shown that the influence of stress on the reaction rate plays the main role. In addition the diffusion coefficient is enhanced significantly for some cases. Quantitative data on these effects were extracted.

ACKNOWLEDGEMENTS

The authors wish to thank B. Brandt, B. Schmiedt and E. Rose for performing the MOS-Technology. Moreover they are indepted to E. Lorenz, R. Schork, A. Daurer and K. Hartlieb for support of TEM microscopy.

REFERENCES

[1] R. B. Marcus, T. T. Sheng, J. Electrochem Soc., June 1982

[2] L. O. Wilson, J. Electrochem. Soc., Vol. 129, No. 4, pp. 831-837, April 1982

[3] T. C. Wu, W. T. Stacy, K. N. Ritz, J. Electrochem Soc., Vol. 130, No. 7, pp. 1563-1566, July 1983

[4] D. B. Kao, Stanford University, Technical Report No. G503-2, June 1986

VERIFICATION OF ION IMPLANTATION MODELS BY MONTE CARLO SIMULATIONS

G. Hobler, S. Selberherr

Institut für Allgemeine Elektrotechnik und Elektronik
Technical University of Vienna
Gußhausstraße 27-29, A-1040 Vienna, AUSTRIA

Monte Carlo simulations are perfectly suited to check the validity of simple models. We investigate 3 models: First, we show that 1D models for the implantation into multilayer targets give reasonable results only if the stopping powers of mask and bulk material are similar. Second, we discuss the construction of 2D point responses from 1D profiles. Third, we show that the method of superposing point responses at mask edges may fail in some cases.

1. INTRODUCTION

The Monte Carlo method is known to be the most powerful tool for the simulation of ion implantation. Analytical models, however, require much less CPU times and allow easy consideration of experimental data. The latter is particularly important because Monte Carlo simulations usually assume amorphous targets so that they do not always yield correct profiles for implantations into crystalline targets [1].

As simple models are usually based on physical considerations and Monte Carlo simulations take physics most accurately into account (apart from the assumption of amorphous targets), Monte Carlo simulations are perfectly suited to check the validity of these simple models. In particular, we will investigate in this paper 1D models for the implantation into multilayer targets (Chapter 2), the construction of 2D point responses from 1D profiles (Chapter 3), and the method of superposing point responses to obtain dopant distributions near mask edges (Chapter 4).

Our Monte Carlo program is, from a physical point of view, similar to the well known program TRIM [2]. One mayor difference of our code is that we evaluate scattering angles by interpolation in a precomputed table. The 2D simulations have been performed with a code which allows arbitrary geometries. Both features are described in Ref. [3].

2. IMPLANTATION INTO MULTILAYER TARGETS

In a recent paper [4], Ryssel discussed 5 models for the implantation into multilayer targets. These models consider 3 situations:

1) Implantation into bare material 1 (concentration profile $C_1(x)$).
2) Implantation into bare material 2 (concentration profile $C_2(x)$).
3) Implantation into a mask/bulk structure with given mask thickness d, where the mask material is material 1 and the bulk material is material 2 (concentration profile $C(x)$).

The purpose of the models is to construct $C(x)$ from $C_1(x)$ and/or $C_2(x)$. $C_1(x)$ and $C_2(x)$ may be obtained by simulations as well as by experiments. The models read:

$$C(x) = \begin{cases} C_1(x) & x < d \\ \alpha \cdot C_2\left(x - d \cdot \left(1 - \frac{R_{p2}}{R_{p1}}\right)\right) & x > d \end{cases} \quad (1)$$

$$C(x) = \begin{cases} C_1(x) & x < d \\ C_2\left(x - \left(d - d'\right)\right) & x > d \end{cases} \quad (2)$$

$$C(x) = \begin{cases} \frac{R_{p2}}{R_{p1}} \cdot C_2\left(\frac{R_{p2}}{R_{p1}} \cdot x\right) & x < d \\ C_2\left(x - d \cdot \left(1 - \frac{R_{p2}}{R_{p1}}\right)\right) & x > d \end{cases} \quad (3)$$

$$C(x) = \begin{cases} \frac{\Delta R_{p2}}{\Delta R_{p1}} \cdot C_2\left(\frac{\Delta R_{p2}}{\Delta R_{p1}} \cdot x\right) & x < d \\ C_2\left(x - d \cdot \left(1 - \frac{\Delta R_{p2}}{\Delta R_{p1}}\right)\right) & x > d \end{cases} \quad (4)$$

$$C(x) = \begin{cases} C_1(x) & x < d \\ \frac{\Delta R_{p1}}{\Delta R_{p2}} \cdot C_1\left(\frac{\Delta R_{p1}}{\Delta R_{p2}} \cdot x - d \cdot \left(\frac{\Delta R_{p1}}{\Delta R_{p2}} - 1\right)\right) & x > d \end{cases} \quad (5)$$

α in (1) and d' in (2) are adjusted in such a way that $\int C(x)dx = \int C_1(x)dx$ ($= \int C_2(x)dx$), what is automatically fulfilled in Models 3, 4, and 5. R_{p1}, R_{p2} denote the mean projected range and ΔR_{p1}, ΔR_{p2} the standard deviation of $C_1(x)$, $C_2(x)$.

Ryssel gave qualitative arguments in favour of Model 1. To investigate the models quantitatively, we have calculated $C_1(x)$, $C_2(x)$ and $C(x)$ by Monte Carlo simulations and then constructed $C(x)$ from $C_1(x)$ and $C_2(x)$ by applying one of the Models 1–5. Comparing the two versions of $C(x)$, one can easily see how good the models are.

Two examples are shown in Fig.1 and Fig.2. Fig.1 shows good agreement between Model 1 and Monte Carlo results for an As-implantation into SiO_2/Si. In Fig.2 can be seen, however, that the model fails completely for a Be- implantation into SiO_2/GaAs. In this case the profile in bare SiO_2 would describe the profile in SiO_2/GaAs much better than the profile constructed by Model 1. This indicates that the models fail, if mask and bulk material have very different stopping powers like SiO_2 and GaAs.

To confirm this result, we have performed simulations for B-, As-, Sb-, and Be-, Si-, Zn- implantations into SiO_2/Si and SiO_2/GaAs, respectively, at 3 different energies and for 3 values of the mask thickness. P-implantations have not been considered because P-profiles in SiO_2 and Si are almost identical. The energies are usually 30 keV, 100 keV, and 500 keV (10, 80, 500 for B and Be), the values for the mask thickness about $\frac{1}{5}R_p$ ("thin"), $\frac{4}{5}R_p$ ("medium"), $\frac{7}{5}R_p$ ("thick"). In order to present the results in a compact manner, we have introduced 4 degrees (cf. Tab.1 and Tab.2): "good" means that the profiles deviate in depth far less than 10%, "fair" means less than 10%, "poor" more than 10%. "catastrophic" has been introduced to indicate that one of $C_1(x)$, $C_2(x)$ would represent the profile in the mask/bulk structure better than $C(x)$ as calculated from the model.

In Tab.1 and Tab.2 there is listed for each mask thickness and each model the number of cases with good, fair, poor, and catastrophic agreement. (Note that the sum of each column is 9, as we have 3 ion species at 3 energies). In Tab.1, which is for SiO_2/Si, it can be seen that the general agreement is quite good, however, only Models 1 and 3 are always "good" or "fair", and Model 1 is slightly better than Model 3, in agreement with Ryssel [4]. On the other hand, all models completely fail for SiO_2/GaAs (Tab.2). Only for thin masks Model 3 gives good results.

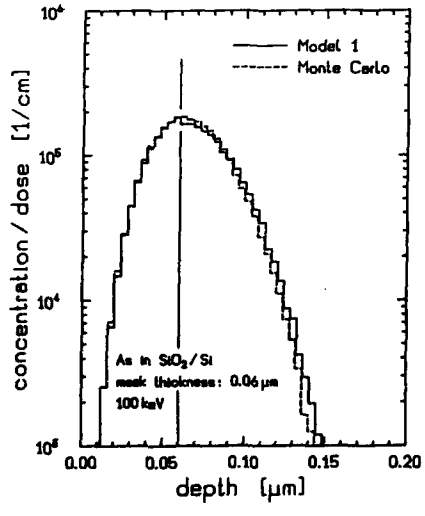


FIGURE 1

As-implantation into Si through a SiO_2 mask.
dashed line: Monte Carlo profile in SiO_2/Si.
full line: Profile in SiO_2/Si due to Model 1, constructed from Monte Carlo profiles in bare SiO_2 and bare Si.

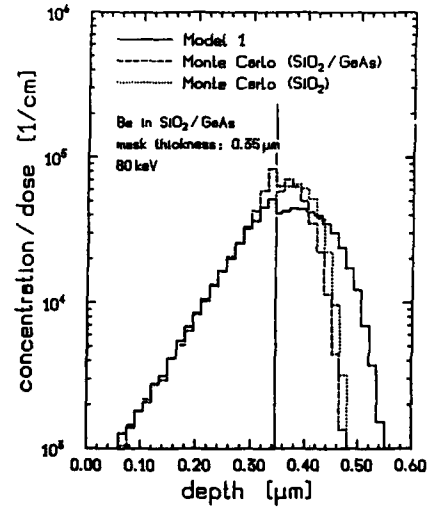


FIGURE 2

Be-implantation into GaAs through a SiO_2 mask.
dashed line: Monte Carlo profile in SiO_2/GaAs.
full line: Profile in SiO_2/GaAs due to Model 1, constructed from Monte Carlo profiles in bare SiO_2 and bare GaAs.
dotted line: Monte Carlo profile in bare SiO_2.

mask	thin					medium					thick				
model	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
good	9	4	9	7	1	6	5	5	3	5	9	9	4	1	9
fair	-	3	-	2	2	3	4	4	1	4	-	-	5	2	-
poor	-	1	-	-	6	-	-	-	5	-	-	-	-	6	-
catastrophic	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-

TABLE 1

Number of cases with good, fair, poor, and catastrophic agreement for implantations into SiO_2/Si.

mask	thin					medium					thick				
model	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
good	3	-	6	2	-	-	-	-	-	-	-	-	-	-	-
fair	2	-	2	-	-	3	-	1	-	-	-	-	-	-	-
poor	-	2	-	3	-	3	1	3	1	1	1	1	1	-	-
catastrophic	4	7	1	4	9	3	8	5	8	8	8	8	8	9	9

TABLE 2

Number of cases with good, fair, poor, and catastrophic agreement for implantations into SiO_2/GaAs.

3. CONSTRUCTION OF POINT RESPONSES FROM 1D PROFILES

Responses to punctiform beams play an important role in the Superposition Method (see Chapter 4). For a long time it was believed that one parameter, namely the lateral standard deviation, would be enough information to construct the 2D point response $C(x,y)$ from the 1D profile $C_{vert}(x)$. This was simply done by multiplying $C_{vert}(x)$ with the lateral Gaussian function $gauss(y)$ given by σ_y:

$$C(x,y) = C_{vert}(x) \cdot gauss(y) \qquad (6)$$

This means that the lateral profile at any depth is a Gaussian function with fixed standard deviation. In a previous paper [5] we have shown that this is not true for Si-targets. The lateral standard deviation depends strongly on the depth, and also the lateral profile is not always well represented by a Gaussian function.

We have now investigated GaAs-targets, and we found quite the same behaviour as for Si: For light ions (Be) the lateral standard deviation decreases with depth (Fig.3) and the lateral kurtosis is smaller than 3. For heavy ions (Zn) the standard deviation increases with depth and the kurtosis may assume large values near the surface. For Si-ions, which lie between the two cases, σ_y does not depend very much on the depth.

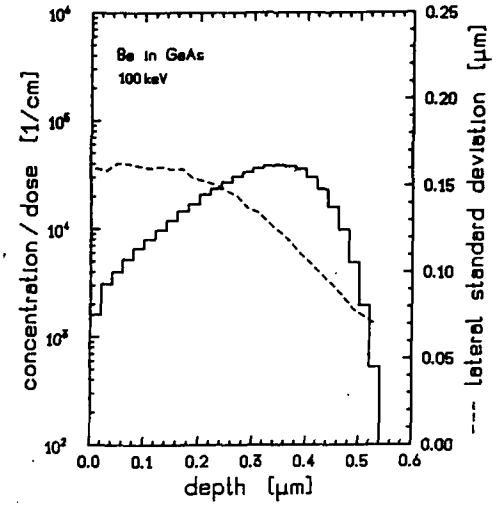


FIGURE 3

Depth dependence of the lateral standard deviation for Be in GaAs (100 keV).

4. SUPERPOSITION METHOD

The superposition law says that the response to a homogenous beam is identical to the sum of responses to punctiform beams which are equidistributed over the width of the homogenous beam. For a rigorous application of this law we would have to know the actual response to every punctiform beam along the surface. In practice, however, point responses are constructed from 1D profiles and may therefore not take into account boundaries other than perpendicular to the beam. In the case of a mask edge those ions are not treated correctly by the superposition method which leave the mask laterally and re-enter the target. The question is now, whether these ions may significantly contribute to the total dopant concentration.

To investigate this question, we have performed Monte Carlo simulations for a simple structure, namely a rectangular mask on a planar bulk. In this case, according to the superposition method, no ions should reach the Si-region which have originally entered the mask. So, if we only expose the mask surface to the computational ion beam, any concentration in the Si-region indicates a failure of the model. We have performed simulations for B- and As-implantations at various energies. The results for B at 100 keV are shown in Fig.4. The concentration in Si is about one order of magnitude lower than the peek concentration of di-

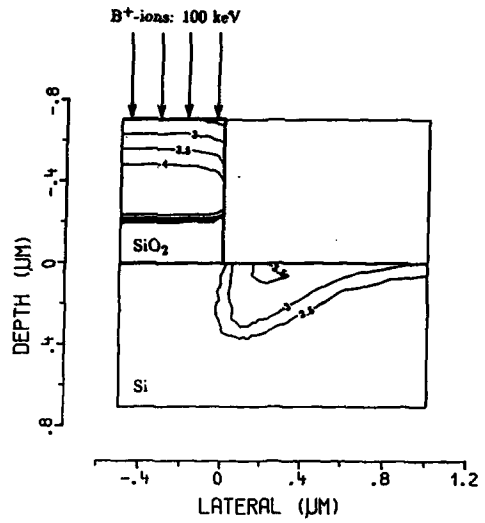


FIGURE 4

B-implantation into Si by a rectangular mask.
The contour lines represent the logarithm of the dopant concentration devided by the dose [1/cm].
Only the mask region is exposed to the beam.

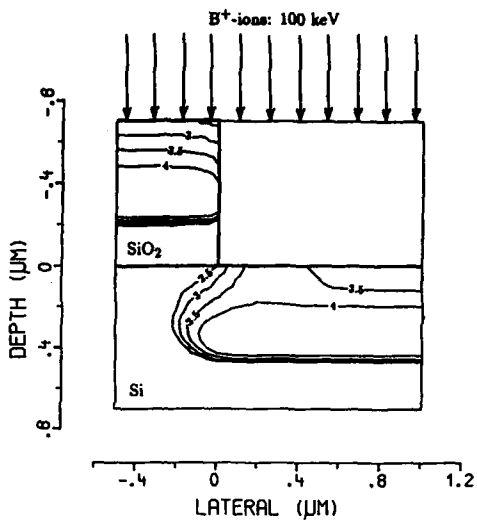


FIGURE 5

B-implantation into Si by a rectangular mask.
The contour lines represent the logarithm of the dopant concentration devided by the dose [1/cm].
The whole simulation area is exposed to the beam.

rectly implanted ions. This is typical for all cases we have simulated.

In Fig.5 it can be seen that this extra concentration —as compared with what is expected by the superposition method—contributes significantly to the total distribution. According to the superposition method, the contour lines labeled by "3.5" and "4" should be straight lines for lateral coordinates from slightly larger than 0 up to 1. Also the contour line labeled by "3" should be seen there.

For As- and low energy B-implantations this extra concentration may be well neglected, because in these cases the profiles have their maximum near the surface and will therefore cover the dopants which have made their way through the mask. A similar situation as in Fig.4 and Fig.5 is expected for high-energy P-implantations.

To avoid this effect, one could use a thicker mask, since the ions which leave the mask laterally will then spread over a wider range. E.g., for a mask thickness of $2\mu m$ in Fig.5 the effect would almost disappear. Another possibility would be to tilt the mask edge. In this case, however, the dopant distribution below the mask edge would be increased.

ACKNOWLEDGEMENTS

This work has been supported by the research laboratories of SIEMENS AG at Munich, by DIGITAL EQUIPMENT CORP. at Hudson, USA, and by the "Fonds zur Förderung der wissenschaftlichen Forschung", project S43/10.

REFERENCES

[1] H. Ryssel, J.P. Biersack:
"Ion Implantation Models for Process Simulation",
In: *Process and Device Modeling* (Ed. W.L. Engl), Elsevier, North Holland, pp. 31–69, 1986.

[2] J.P. Biersack, L.G. Haggmark:
"A Monte Carlo Computer Program for the Transport of Energetic Ions in Amorphous Targets"
Nucl. Instr. Meth., Vol. 174, pp. 257–269, 1980.

[3] G. Hobler, S. Selberherr:
"Efficient Two-Dimensional Monte Carlo Simulation of Ion implantation"
Proc. NASECODE V Conf., Dublin, 1987.

[4] H. Ryssel, J. Lorenz, K. Hoffmann:
"Models for Implantation into Multilayer Targets"
Appl. Phys., Vol. A41, pp. 201–207, 1986.

[5] G. Hobler, E. Langer, S. Selberherr:
"Two-Dimensional Modeling of Ion Implantation with Spatial Moments"
Sol.-State Electron., Vol. 30, No. 4, pp. 445–455, 1987.

A SIMPLIFIED MODEL FOR THE CHARACTERIZATION OF ANTIMONY ION IMPLANTATION AND DIFFUSION ON SILICON

REUSI INÊS FONSECA

Laboratório de Microeletrônica da Escola Politécnica da USP
Departamento de Engenharia de Eletricidade, P.O. Box 8174
05508 São Paulo, Brasil

The main purpose of this work is to show how with a simple analytical model of diffusion, using only $R_\square$ and x_j experimental data, is it possible to calculate the relationship between the electrically active charge and the initial implanted dose of Antimony in Silicon. It will be shown that $Q_{el}/Q_{dose} < 1$ in agreement with many others authors[1,2,3], and a lower diffusion coefficient than that commonly used by SUPREM II simulator is achieved through this model, which allows good fit with the experimental data.

1. INTRODUCTION

High dose implantations of Antimony in Silicon studies are receiving considerable attention in recent years, by many authors[1,2,3,4], owing to their applications as an impurity source in the fabrication of buried-layers in high speed, low power dissipation Bipolar Transistors[5,6], and resistors with special characteristics[7], convenient for VLSI circuits. It is well known that for small concentrations of Sb in Si, almost 100% substitutionallity and electrical activity is achieved[2,3], whereas high concentrations exceeding the solid solubility limits of Sb in Si form metastable solutions and cause segregation effects, indicating that a fraction of substitutional Sb is electrically inactive. A complete characterization of these high concentrations of Sb in Si, their decomposition into precipitates, informations about the exact Sb atoms crystallographic location in the lattice and the degree of their electrical activity has been extensively studied, through a variety of analytical techniques[2,3].

This work shows through a simplified analytical model of diffusion, using only $R_\square$ and x_j experimental data, that is it possible to calculate the relationship between the electrically active charge, Q_{el}, and the initial implanted dose, Q_{dose}, of Sb in Si, together with the following parameters: diffusion coefficient D, mean mobility $\bar{\mu}$ and carrier concentration $\bar{C}$.

These results are in good agreement with those published earlier by many others authors, indicating that this model, although very simple, is sufficiently adequated to describe the fenomena correlated with the Sb diffusion on Si.

2. THE ANALYTICAL MODEL

After annealing and drive-in diffusion, impurity concentration redistribution from an ion-implanted source can be treated as a Gaussian profile, with fixed amount of impurities, given by Q_{dose}. Therefore, it is well known that only a fraction of these Q_{dose} is electrically active, named Q_{el}, and given by:

$$Q_{el} = \frac{1}{q\,\bar{\mu}\,R_\square} \qquad (1)$$

From $R_\square$ experimental data, the only way to calculate the actual Q_{el} is through the mean mobility $\bar{\mu}$ variations, which have a dependence over the mean impurity concentration $\bar{C}$. These parameters can be interconnected through the following definitions:

$$\bar{C} = \frac{Q_{el}}{x_j} \qquad (2)$$

and

$$\bar{\mu} = \mu min + \frac{\mu max - \mu min}{1 + \left(\frac{\bar{C}}{N_{ref}}\right)^{\alpha}} \qquad (3)$$

where:

$\mu min = 86{,}5\ cm^2\ V^{-1}\ s^{-1}$

$\mu max = 1354{,}5\ cm^2\ V^{-1}\ s^{-1}$

$\bar{C}$ = mean impurity concentration

$\alpha = 0{,}91$

$N_{ref} = 1{,}3 \times 10^{17}\ cm^{-3}$

It can be shown that for a long time diffusion the junction depth, x_j, is given by:

$$x_j^2 = 4Dt\ ln\ \frac{Q_{dose}}{C_B \sqrt{\pi D_i}} \qquad (4)$$

A linear function is obtained from $x_j \times \sqrt{t}$ plotting, whose gradient $dx_j/d\sqrt{t}$ allows the diffusion coefficient D determination from relation (4):

$$\sqrt{D} = \frac{dx_j}{d\sqrt{t}} \cdot \frac{1}{2\left[ln\ \frac{Q_{dose}}{C_B \sqrt{\pi D_i}}\right]^{1/2}} \qquad (5)$$

The first step in an iterative calculation between the relations (1), (2), (3) and (5), considers that all implanted impurity is electrically active, so that it is obtained the first values of D, $\bar{C}$, $\bar{\mu}$, and Q_{el}. Subsequent iteration calculations will allow to get the actual values of $\bar{D}$, $\bar{C}$, $\bar{\mu}$ and Q_{el}, when convergence is achieved.

As will be seen in the next item, using only the x_j and $R_\square$ experimental data it will be possible to obtain results comparable with those from another authors, whose experimental data were obtained from several technological facilities, like RBS analysis, Mössbauer spectroscopy, Hall-effect measurements, and so on[2,3].

3. EXPERIMENTAL PROCEDURES

All Sb implantation were performed in a homemade equipment, with an energy E = 100 Kev and dose $\phi = 5 \times 10^{15} cm^{-2}$, at room temperature, in a 7° off-axis direction, into silicon wafers type P, <100>, ρ = 10-20Ω.cm. After typical annealing at low temperature, T = 500°C, and O_2 ambient, for 60min, it was carried out the dopant diffusion in O_2 ambient, T = 1200°C, in 4 different times, 4, 9, 16 and 25 hours. Subsequent measurements of $R_\square$ and x_j of those samples were used to calculate the diffusion coeeficient, mean impurity concentration, mean mobility and electrically active charge by means of that iterative procedure suggested by the model described earlier in this paper.

4. RESULTS AND DISCUSSION

Figure 1 shows experimentally determined junction dephts, x_j, and sheet resistance mean values $R_\square$, obtained for several samples over a wide range of heat treatment times, compared with the simulated values produced by SUPREM II, only with the purpose of a more detailed overview about Sb diffusion on Si behavior:

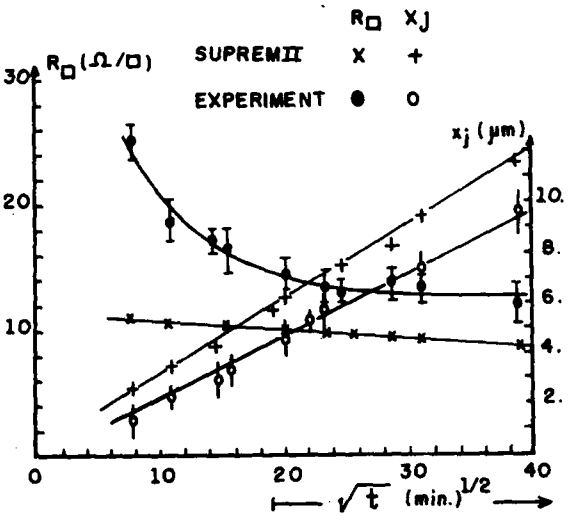


FIGURE 1

Comparison between SUPREM II and experimental $R_\square$ and x_j values.

As SUPREM II has not an Antimony cluster model, the predict data are overestimated in the sense that all implanted impurity is considered as electrically active, and the experimental results are below those simulated; moreover, the

saturation on $R_\square$ experimental data indicates a saturation on Q_{el}, due to the formation of extended defects like new Sb clusters or precipitates during the heat treatment[2], not considered by SUPREM II.

From $(dx_j/d\sqrt{t})_{EXP}$ of figure 1 and using the analytical model proposed on this work, it was calculated the new diffusion coefficient D value ($D \cong 1{,}16 \times 10^{-11} cm^2.min^{-1}$), and consequently, the $\overline{C}$, $\overline{\mu}$ and Q_{el} parameters. Modyfying SUPREM II by introducing the new values of D and Q_{el} as the initial implanted dose, it was obtained a good fit with experimental $R_\square$ and x_j values, as illustrates figure 2 below:

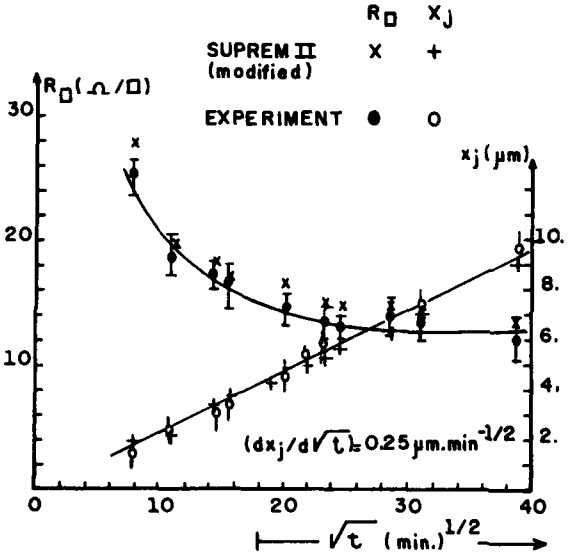


FIGURE 2

Comparison between modified SUPREM II and $R_\square$ and x_j experimental values.

Finally, on figure 3 it is shown the relation $Q_{el}/Q_{dose} \times \sqrt{t}$, where two points must be emphasized:

- $Q_{el}/Q_{dose} < 1$ indicates that only a fraction of implanted dose is electrically active, probably that one on undisturbed substitutional lattice sites. This percentual result shown in figure 3 is in good agreement with others authors[1,2,3];
- the saturation of Q_{el}/Q_{dose} for long time diffusion is an expected result, and indicates a reduction of the electrically active fraction of Antimony, probably by the formation of precipitates[2] or Sb-vacancy complexes[2,4].

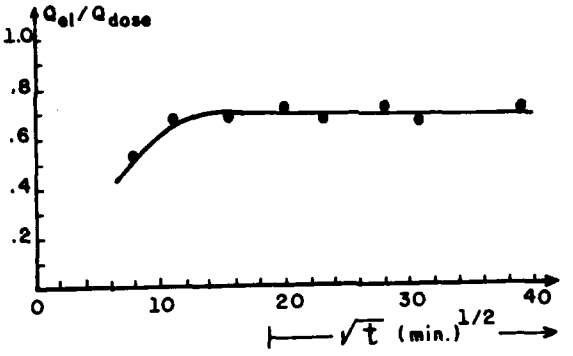


FIGURE 3

Q_{el}/Q_{dose} as calculed by the model proposed on this work.

Although not presented here the mean concentration, $\overline{C}$, and mean mobility, $\overline{\mu}$, as calculated by the model[6], seems to be in agreement with results from another authors[2,3,8].

5. CONCLUSIONS

In this work it was presented a very simple analytical model, that allows to get informations about the electrically active fraction of Antimony implanted on Silicon, taking into account only $R_\square$ and x_j experimental data. The results of the model have shown that the diffusion coefficient of Antimony on Silicon is lower than that used by SUPREM II simulator, and that only a fraction of the initial implanted dose is electrically active, as previously published by many others authors. The modification of SUPREM II by introducing these new data allowed a good fit with experimental $R_\square$ and x_j data.

ACKNOWLEDGEMENTS

The author would like to thank the students H. Peres and E. Galeazzo, for making the ion implantation and the experimental measurements, respectively, and Mrs. M. Brito for typing this manuscript. Helpful suggestions of Dr. S.Solmi are also acknowledged.

The partial financial support by FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo,is gratefully acknowledged by the author.

REFERENCES

1. JOSQUIN, W.J.M.J. and TAMMINGA, Y., Applied Physics, 15, (1978), 73-78.

2. NYLANDSTED LARSEN, A., et al, MRS Europe, (1985), 319-324.

3. NYLANDSTED LARSEN, A., et al, J. Appl. Phys., 59 (6), (1986), 1910-1917.

4. FAIR, R.B., et al, J. Mater. Res., 1 (5), (1986), 705-711.

5. TANG, D.D., et al, IEEE Trans. Electron.Dev., 27 (8), (1980), 1379-1384.

6. FONSECA, R.I., Ph.D. Thesis, Escola Politécnica, USP, (1985).

7. KU, S.M., and CHU, W.K., Solid-State Electronics, 22, (1979), 719-722.

8. JOHANSSON, N.G.E., and MAYER, J.W., Solid-State Electronics, 13 (1970), 123-130.

GLASS REFLOW MODELING FOR PROCESS OPTIMIZATION

A. TISSIER, A. PONCET and J.F. TEISSIER
CNET-Grenoble - France

1 INTRODUCTION

PSG and BPSG are intensively used in VLSI processes for their flow capability . In a micronic multilevel metallisation technology, it is necessary to control the flow annealing which tends to smooth the topology, particularly in two places : the gate overlap and the contact window steps . In the literature, work has been mentionned which deals essentially with measurements on SEM views of the tangential angle of the layer at the step edge as a funtion of the annealing parameters and the glass composition [1],[5],[7]. A new approach is presented here, which combines experimental results with numerical simulations of glass reflow, in order to predict the "optimal" annealing, i.e. an increased planarity and a minimization of parasitic thermally activated phenomena (dopant diffusion). Coupling SEM measurements and numerical simulations allows to process only one test pattern and, furthermore, to extrapolate the results to any case.

2 PHENOMENOLOGICAL STUDY OF THE VISCOUS FLOW

In order to minimize the induced technological dispersions, the simplest test pattern is chosen, i.e. a rectangular glass slab (0.8 µm height and 4 µm width). After cleaving , the geometrical evolution of cross sections is studied by SEM measurements as a fonction of the RTA parameters (T from 950 o C to 1190 o C , t from 10 to 80 s) and the glass composition (6% w/o P to 8.9% w/o P PSG and 5% w/o B, 5% w/o P BPSG). As depicted on figure 1 the glass reflow leads to the modification of the following three geometrical parameters

- the angle, θ,
- the thickness at the middle of the step, h,
- the curvature radius, R;

The first parameter being the more sensible to the reflow annealing, it is chosen to quantify the viscous deformation.

Experiments made on both gate overlap and test pattern show that for θ equal to 15 degrees, the planarisation is acceptable. We observe that even in simplest cases, it is difficult to directly compare experimental and simulated profiles because of the dispersion, wafer to wafer or run to run, on data related to the slab formation, i.e. thickness, CD, angle after etching and local glass composition, therefore, it is necessary to average data.

The evolution of θ according to the time is measured for different temperatures , and for various glasses. Let t denote the annealing time which leads to θ=15 degrees. t values are extracted from these measurements; next, ln(t) is plotted as a function of 1/T (figure 2).

From these curves it can be observed that :

- the linearity of this function allows to fit physical parameters related to glass viscosity by carriing out linear regressions (see section 3);
- moreover, when these curves are superimposed with similar curves related to other thermally activated phenomenon (here boron diffusion) it is possible to identify the optimal temperature range for a given glass.

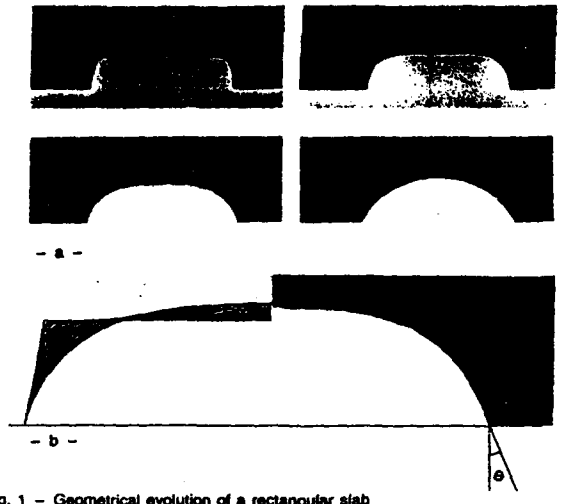


Fig. 1 – Geometrical evolution of a rectangular slab
– a – SEM views (8.6% w/o P PSG, T = 1190 C, t = 5, 10, 20 and 40 s)
– b – Simulation results compared with SEM view at t = 20 s.

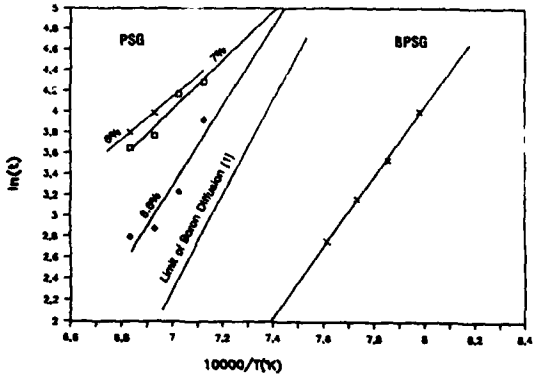


Fig. 2 – Logarithm of the time necessary to get Θ = 15° as a function of 1/T

3 NUMERICAL SIMULATION OF GLASS REFLOW

3.1 Viscous Flow Model

Under usual processing temperatures, the viscosity of passivation glasses is sufficiently high to assume that stationnary Stokes equations are valid for modeling their reflow. The driving force is the surface tension [8], which tends to smooth the free surface by increasing curvature radius in such a way that the cross section of any bounded slab of glass tends to be a piece of perfect disk (figure 1). However, such final shapes have no practical interest and are not valid for unbounded slabs, moreover, intermediate stages can not be accurately predicted under geometrical considerations only [4]. Therefore, computer simulations are necessary, i.e. discretization of Stokes equations. These equations can be summarized as follows in the 2-D case:

$$(3.1) \qquad v\ \mathrm{Div}\ (\ \mathrm{Grad}\ \vec{V}\ ) = -\ \mathrm{Grad}\ p$$

and,

$$(3.2) \qquad \mathrm{Div}\ \vec{V} = 0 \qquad \text{in the material,}$$

where $\vec{V}=(Vx,Vy)$ is the local velocity, v is the viscosity and p the internal pressure;

(3.3) $Vx = Vy = 0$ (non-slip condition) on the interface between glass and substrate,

(3.4) $Vx = 0$ (slip condition) along symetry axes and lateral sections,

(3.5) $p = Y / R$ along the free surface, where Y is a surface tension coefficient and R is the curvature radius.

A major application of glass reflow simulation concerns contact holes; for that purpose, an axisymetric expression of the equations has been set under a variational form; 3-D effects have been clearly evidenced in numerical experiments: Θ becomes much smaller when the radius of a contact hole decreases (figure 3).

3.2 Viscosity Fitting

The main advantage of the above model is its linearity according to v/Y; this ratio can be easily identified from experiments as follows:

1. arbitrary v/Y ratio and time scale are chosen (let say v/Y=1), then computer simulation is performed;

2. numerical results are compared with measurements (figure 1) in order to set the time scale which corresponds to a given glass: let tn be the time necessary to reach a given D value in the computer simulation, and tm the corresponding time deduced from experiments; therefore, the actual value of v/Y is tm/tn.

According to the two linearities mentionned above, relation (2.1) can be re-written, first:

$$(3.6)\quad \ln(t) = A + B/T$$

and then

$$(3.7)\quad \ln(v/Y) = A + B/T - \ln(tn)$$

which confirms the classical expression [2]:

$$(3.8)\quad v/Y = u0 \,.\, \exp(E/kT)$$

while giving an straightforward evaluation of u0 and E parameters:

$$(3.9)\quad u0 = \exp(A)/tn$$

$$(3.10)\quad E = k \,.\, B$$ where k is the Boltzmann constant.

By using measurements depicted on figure 1 for a 4 μm long and 0.8 μm high step, this method leads to values for E and u0 which are presented on table 1; however, the reproducibility of the slab dimensions and of RTA parameters, the accuracy of measurements and the temperature range are too low to quantify the dependence of E and u0 versus glass composition.

3.3 Numerical Schemes And Computer Environment.

Equations (3.1)-(3.5) are discretized by using classical 3-node triangular finite elements, meshes are automatically generated and refreshed, in the same way as in LOCOS simulation [7]. Incompressibility condition (3.2) is taken into account iteratively, by using classical Uzawa algorithm. Surface condition (3.5) is expressed through a boundary integral in the variationnal form of the equations.

Impurity diffusion and glass reflow have been coupled in TITAN process simulator [3], in order to achieve technological parameter optimization which has been mentionned in Section 1.

4 CONCLUSION

A simple linear viscous flow model has been presented in order to predict PSG or BPSG glass reflow. A method has been presented for identifiing viscosity parameters for any given glass, in order to optimize glass reflow, anywhere on the wafer.

However, the application of this approach depends drastically on the initial structure (composition, shape of the slab,...) and on RTA parameters variations; therefore, a general expression of viscosity versus temperature and glass composition cannot be set as long as these data are not accurate enough.

	PSG			BPSG
	6%	7%	8.6%	
E (eV)	5.15	4.56	1.25	1.08
u0	10^{-13}	10^{-11}	10^{-29}	10^{-28}

TABLE 1. Viscosity parameters from fits between measurements (Fig. 1) and computer simulation

5 REFERENCES

[1] N.S. ALVI and D.L. KWONG "Reflow of PSG by Rapid Thermal Annealing" Symp. on reduced temperature processing for VLSI, Proceedings Vol 86-5, edited by the Electrochemical Soc. (1986).

[2] D. CHIN "Two-Dimensional Oxidation, Modeling and Applications" PhD. Stanford, June 1983

[3] A. GERODOLLE, S. MARTIN and A. MARROCCO, "Finite Element Method Applied to 2-D MOS Process Simulation and defect diffusion: Program TITAN", NASECODE IV Conf. Proceedings, Boole Press, June 1985.

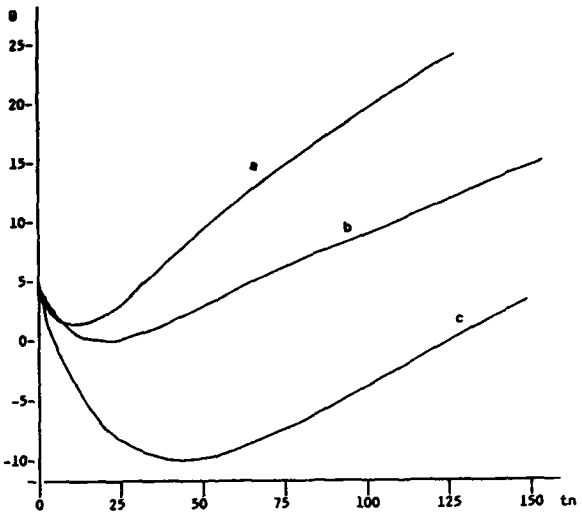
[4] R.A. LEVY and K. NASSAU " Reflow Mechanism of Contact Vias in VLSI Processsing" J. Electrochem. Soc., Vol. 133 No. 7, pp. 1417-1424 (1986).

[5] J.S. MERCIER, R.P. BEERKERNS, I.D. CALDER and H.M. NAGUIB, Electrochemical Soc. ext. abstracts No. 420 Vol. 84-2 p. 607 (1984).

[6] A. PONCET, "Finite Element Simulation of Local Oxidation of Silicon", IEEE Trans. on Computer-Aided Design, Vol. CAD-4 No.1, pp. 41-53, 1985.

[7] T.O. SEDGWICK, F.M. D'HEURLE and S.A. COHEN, J. Electrochem. Soc. Vol. 131, p. 2446 (1984).

[8] P. SUTARDJA, Y. SHACHAM-DIAMAND and W.G. OLDHAM, "Two-Dimensional Simulation of Glass Reflow and Silicon Oxidation" I.E.D.M. Conf. , Los Angeles, Dec. 1986.



igure 3. θ-angle versus time:
-a- 0.4 μm thick and 4 um long slab,
-b- 0.8 μm thick and 4 um long slab,
-c- cylindrical contact hole,
hight=0.4 um, radius=0.25 um.

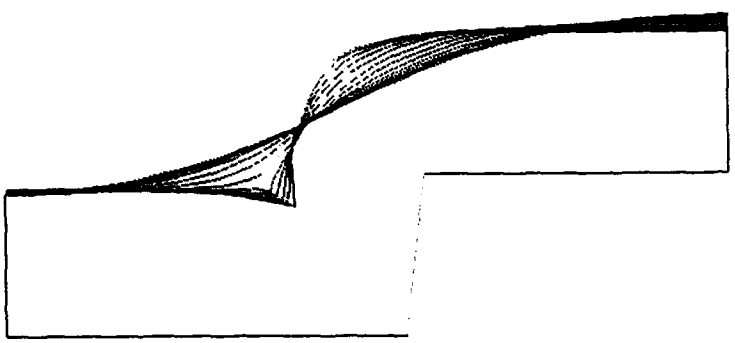


Figure 4. Numerical simulation of glass reflow at gate overlap

MONTE-CARLO ION IMPLANTATION AND COMPOSITE

A.Barthel, J.Lorenz, H.Ryssel*

Fraunhofer-Arbeitsgruppe für Integrierte Schaltungen,
Artilleriestrasse 12, D-8520 Erlangen, Germany

Analytical methods for the description of ion implantation show good agreement with experiment and Monte-Carlo simulations in most cases. Problems arise with special geometries such as trenches. To be able to simulate implantation and diffusion in such cases, a Monte-Carlo interface has been added to the process simulation program COMPOSITE.

1. INTRODUCTION

To meet the needs of shrinking device dimensions, process simulation programs are required which use accurate physical models for the simulation of process steps, use efficient algorithms to reduce computing time and are able as well to deal with a complete process sequence as to transfer the results as input for device simulation. These three requirements are very hard to be fulfilled with one simulation tool, as accurate process models very often require large computing time, for instance in case of Boltzmann transport equation calculations or Monte-Carlo simulations [1].

In the following, the approach to ion implantation used in COMPOSITE [2] is briefly mentioned along with its limitations. The Monte-Carlo interface which has been added to COMPOSITE is described and its application is shown.

2. COMPOSITE

The universal two-dimensional process simulation program COMPOSITE (Complete Modeling Program of Silicon Technology) is a user-friendly and easily-portable tool for the simulation of ion implantation, diffusion, oxidation, etching, lithography and layer deposition. For the simulation of ion implantation, analytical equations are used for the dopant concentration profiles. This includes the well-known Pearson IV-distributions [3] along with range parameters from experiments for the vertical dopant concentration profile in one layer, a lateral convolution with a Gaussian profile and a special multilayer model [4], which takes into account the different stopping powers of the layers. In figure 1, results obtained with COMPOSITE for an implantation of 60 keV phosphorus at an Al_2O_3-mask edge are compared to results obtained using the widely used Runge model [5] (broken lines), which assumes the same stopping power for all layers. According to the Runge model, the Al_2O_3 layer would not be thick enough to mask the silicon. From the COMPOSITE-result it can be seen that the Al_2O_3 thickness is sufficient to stop the ions.

3. MONTE-CARLO SIMULATIONS

Other methods for the simulation of ion implantation such as Monte-Carlo simulations [1] require much more computing time in comparison to analytical models and are, therefore, not suited for permanent use in a process simulation tool. But they are very important for the evaluation of analytical

*also: Lehrstuhl für Elektronische Bauelemente,
Universität Erlangen-Nürnberg,
Artilleriestrasse 12, D-8520 Erlangen, Germany

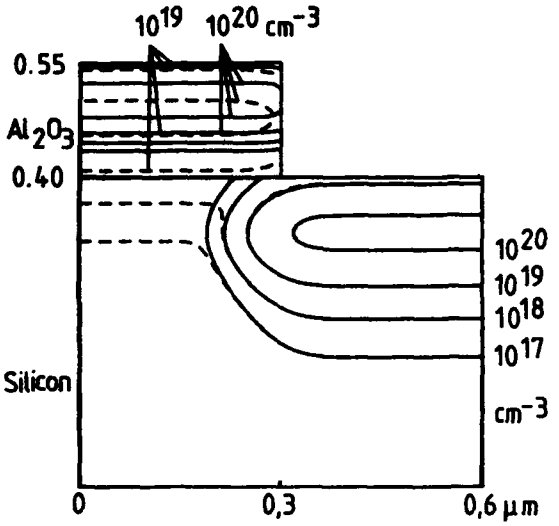


FIGURE 1

COMPOSITE-simulation of implantation of a 10^{15} cm^{-3} dose of phosphorus at an energy of 60 keV near an Al_2O_3 mask edge.

descriptions of implantation profiles and for the simulation of implantations into geometries which cannot be described adequately by analytical models. One main point of interest is the implantation into trenches in silicon.

Figure 2 shows the result from a Monte-Carlo simulation with a modification of TRIM, TRIMSURF [6], of an implantation of 150 keV phosphorus into a 2 µm deep and 0.4 µm wide silicon trench. 200 000 particles were used for this simulation. In this example, a sidewall-doping by ions which have been scattered out of one sidewall and have been re-implanted into the other sidewall can be seen. This sidewall doping is of great importance and cannot be accessed by analytical models. Therefore, it is very important to use such Monte-Carlo results within general simulation tools.

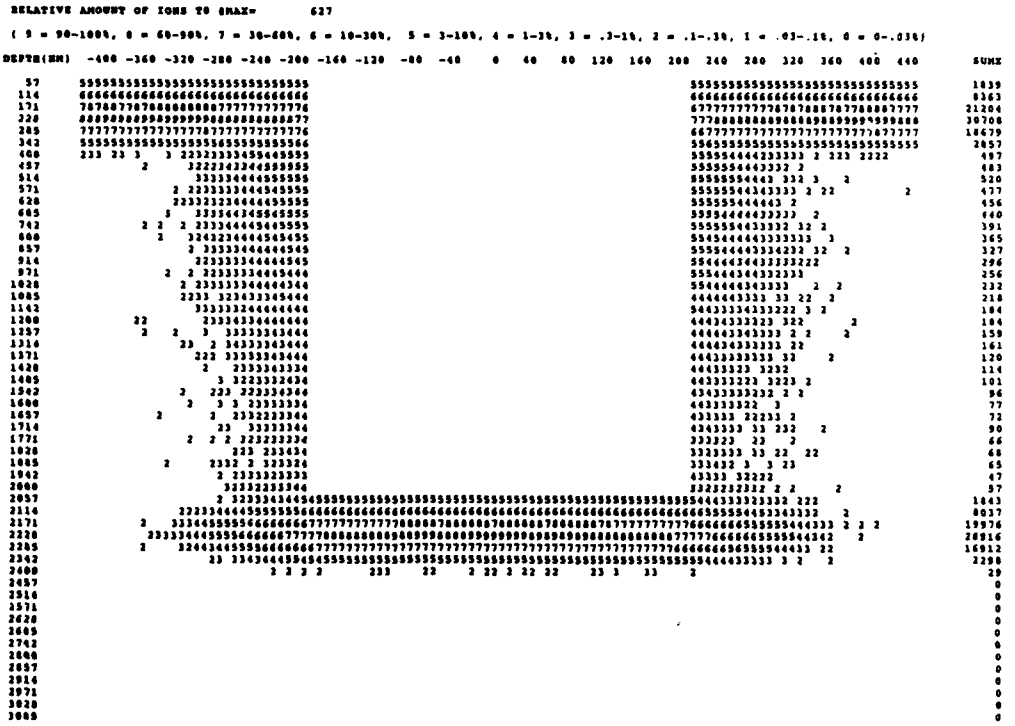


FIGURE 2

Monte-Carlo simulation of implantation of phosphorus at an energy of 150 keV into a 2µm deep and 0.4 µm wide silicon trench. 200 000 particles were used for this simulation.

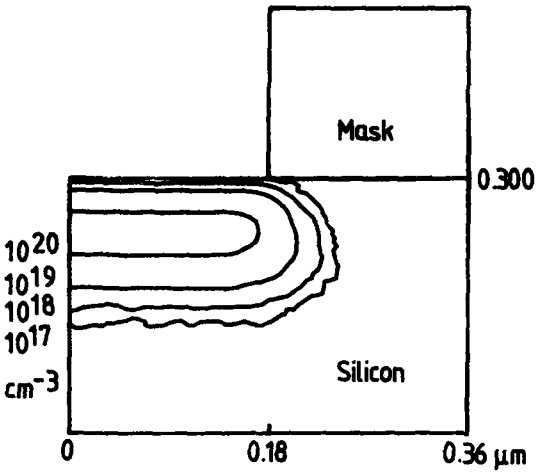


FIGURE 3

Monte-Carlo simulation of implantation of arsenic at an energy of 100 keV into silicon near a mask edge.

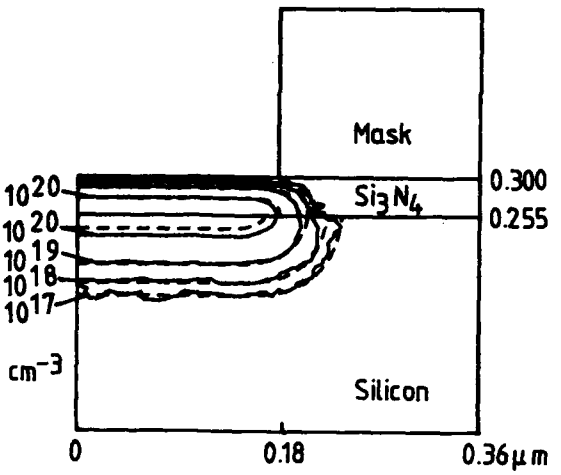


FIGURE 4

Comparison of TRIMSURF and COMPOSITE results for implantation of a dose of 10^{15} cm^{-3} arsenic at an energy of 100 keV into a two-layer structure near a mask edge.
Drawn line: TRIMSURF, broken line: COMPOSITE

4. USE OF MONTE CARLO SIMULATIONS FOR COMPOSITE

To be able to transfer results from Monte-Carlo simulations to COMPOSITE, an interface has been implemented.

First, some modifications to TRIMSURF have been done. These includes the gathering of the particle distribution data in a COMPOSITE-compatible shape. Furthermore, smoothing by neighborhood averaging is done to reduce statistic fluctuations with the data: Concentrations are recalculated as arithmetic means of the point in question and its eight nearest neighbors. The dopant concentration arrays are then stored to a file.

Second, COMPOSITE reads these data from the file and scales them according to the implantation dose desired.

In figure 3, an example for a Monte-Carlo simulation of an ion implantation of 100 keV arsenic into a silicon layer near a mask edge is shown. The Monte-Carlo data have been transferred to COMPOSITE. In the equiconcentration line plot the fluctuations in the third and fourth contour line result from the limited number of particles used with the Monte Carlo calculations. The maximum of the lateral spread of the ions implanted does not coincide with the maximum of the vertical distribution. This indicates the depth dependence of the lateral straggling, studied in earlier publications [7,8,9]. Equations for this depth dependence have been proposed [8], but they presently cannot be used for the simulation of ion implantation in crystalline silicon, because they need not only vertical moments but also lateral and mixed range moments, in total 8 parameters. The lateral kurtosis and the two mixed moments requested have not yet been measured or calculated for crystalline silicon. Therefore, the depth dependence is not included in COMPOSITE. For amorphous silicon, this model shows good agreement with Monte-Carlo simulation [8].

In contrast to figure 3, the silicon is covered by 45 nm Si_3N_4 in the example shown in figure 4. This is done to show the influence of a thin layer on the lateral spread in the silicon substrate. Since the lateral straggling in the nitride is smaller than in silicon because of the higher density of nitride, the

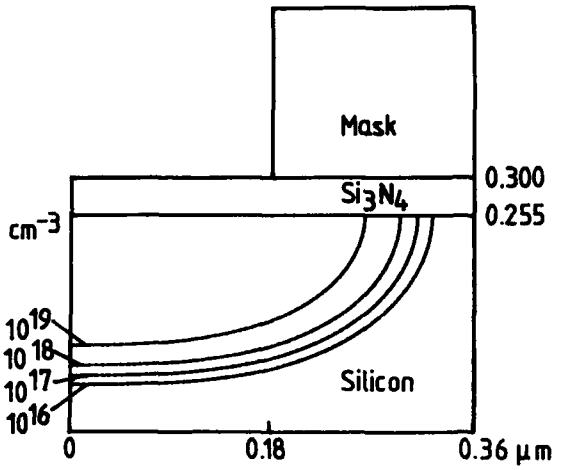


FIGURE 5

30 min diffusion at 1000 °C of the Monte-Carlo result shown in figure 4.

lateral spread of the implantation profile in the silicon is reduced in comparison to figure 3. Therefore, a multilayer model for the depth dependency of the lateral spread is necessary.

The broken lines in figure 4 show the corresponding COMPOSITE-results. Here, a constant lateral straggling was used within one material. Therefore, a discontinuity of the lateral spread is predicted by COMPOSITE, but TRIMSURF shows a nearly continuous behavior of the lateral spread at the interface. For the vertical dopant distribution, the agreement between COMPOSITE and Monte-Carlo is very good, except close to the silicon/nitride interface. The discontinuity of the vertical distribution at the interface results from the lower stopping power of the silicon: Therefore, less particles come to rest below the interface. This effect is less pronounced with the Monte Carlo data because of particles backscattered.

This example shows that as long as no full set of eight range parameters for the materials involved, including crystalline silicon, is available, no accurate simulation of the lateral spread is possible with analytical models. If the differences between amorphous and crystalline range parameters can be neglected, it is worthwhile to transfer Monte-Carlo data to COMPOSITE to perform the simulation of further process steps.

In figure 5, the COMPOSITE result of a 30 min diffusion at 1000°C of the doping profile from figure 4 is shown. The statistic fluctuations present in figure 4 have been removed by the diffusion.

5. CONCLUSION

Though the analytical equations for ion implantation used in COMPOSITE are able to describe dopant profiles adequately in most cases, the simulation of important effects such as trench implantation and depth dependence of the lateral spread presently needs time-consuming Monte-Carlo calculations. The Monte-Carlo interface implemented in COMPOSITE allows now for introducing Monte-Carlo data into a process sequence.

REFERENCES

[1] Biersack, J.P. and Haggmark, L.G., Nucl.Instrum. Methods 174 (1980) 257

[2] Lorenz, J., Pelka, J., Ryssel, H., Sachs, A., Seidl, A. and Svoboda, M., IEEE Trans. El. Dev. ED-32 (1985) 1977

[3] Hofker, W.K., Philips Res. Repts., Suppl. No. 8 (1975)

[4] Ryssel, H., Lorenz, J. and Hoffmann, K., Appl. Phys. A 41 (1986) 201

[5] Runge, H., Phys. Stat. Sol. (A) 39 (1977) 595

[6] Ryssel, H., Lorenz, J. and Krüger, W., Nucl. Instrum. Methods B 19/20 (1987) 45

[7] Hobler, G., Langer, E. and Selberherr, S., Two-dimensional modelling of Ion-Implantation, in: Board, K., Owen, D.R.J. (ed.), Simulation of Semiconductor Devices and Processes, Vol 2, (Pineridge Press, Swansea, U.K., 1986) pp. 256-270

[8] H. Ryssel et al., 4th report on "Zweidimensionale Prozeßsimulation vollständiger technologischer Prozeßabläufe", Erlangen, 1986

[9] Ashworth, D.G. and Oven, R., Computer simulation of the lateral spreading of implanted ions, ESSDERC '86

P2.1.10

EQUILIBRIUM SOLUBILITY OF ARSENIC AND ANTIMONY IN SILICON

R.ANGELUCCI, A.ARMIGLIATO, E.LANDI, D.NOBILI, S.SOLMI

CNR - Istituto LAMEL, Via Castagnoli 1, 40126 Bologna, Italy

Equilibrium solid solubility of arsenic and antimony in silicon is derived by Hall and resistivity measurements after suitable annealing. For both elements, the solubility shows a linear trend versus reciprocal temperature.

1. INTRODUCTION

The knowledge of solid solubility of dopants in silicon is essential for a correct process simulation and is important for basic understanding. In the case of Group V dopants largely scattered values are reported in literature for arsenic and antimony. Moreover, the knowledge is even poorer in the range 700÷900°C, which is of high interest in the future VLSI-ULSI processing.

A research activity on the solubility and precipitation of silicon dopants is performed since several years at LAMEL Institute. Particular emphasis was given to the study of electrically inactive phosphorus and arsenic; an assessment of this problem was attempted by Nobili a few years ago /1/.

This paper reports the results of accurate equilibrium carrier density determinations as a function of temperature, performed on polysilicon films heavily doped with antimony and arsenic by ion implantation. The carrier density was determined after annealing at increasing temperatures, a time consuming procedure which, on the other side, is most suitable to accomplish with the equilibrium conditions.

2. EXPERIMENTAL

Poly-silicon films were deposited in a chemical vapour reactor at 660°C, onto previously oxidized single crystal wafers. The film thickness (0.45 μm) was accurately determined by a Taylor Hobson Talystep.

For Sb doped films the implantation energy and dose were 150 keV and 2.1×10^{16} at/cm^2 respectively, while As was implanted at the energy of 100 keV, and dose 4.0×10^{16} at/cm^2. Each composition was then separately heated 3 hours at 1100°C to recover the damage and redistribute the dopant. Specimens were successively annealed for 1000 h at 600°C, then at temperatures incresing in steps of 25°C up to 900°C and subsequently in steps of 50°C up to 1300°C. To avoid out diffusion processes the first high temperature heat treatment was performed in a slightly oxidizing atmosphere (90% nitrogen + 10% oxygen).

The carrier density and mobility were determined by Hall effect and sheet resistivity measurements, using the Van der Pauw geometry defined with a photolitographic process. The same techniques, alternated with stripping of silicon by anodic oxidation and etching, were used for carrier profile measurements. The average grain size of the polysilicon films, after the high temperature annealing (1100°C, 3h) was checked by transmission electron microscopy observations. The obtained values were 0.9 μm and 0.2 μm for As and Sb, respectively.

3. RESULTS AND DISCUSSION

The carrier density plot vs reciprocal tempe-

rature for Sb doped specimens, which is reported in Fig.1, shows an initial decrease which can be attributed, as it is discussed below, to the formation of the conjugate liquid phase. A minimum is attained at 800°C followed by dissolution which takes place with increasing temperature. Equilibrium values of the carrier density n_e are obtained in the dissolution stage, above 850°C, after a transient which is due to the size effect. The equilibrium values of n_e in the temperature range 850-1150°C follow very tightly the law:

$$n_e = 3.8 \times 10^{21} \exp(-0.56 \text{ eV}/kT) \qquad \text{cm}^{-3}$$

Above 1150°C the experimental n_e values show a deviation from this trend. This was expected as in the Si-Sb equilibrium diagram the conjugate liquid phase undergoes a drastic reduction of the content of antimony /2/.

The results obtained on As doped specimens are shown in Fig.2, which reports the carrier concentration as a function of reciprocal temperature. In this case, due to the higher diffusivity of arsenic with respect to antimony, the minimum was attained at a lower temperature, about 650°C, and equilibrium values of the active dopant were obtained for $T \geq 700$°C. In the temperature range 700-900°C the corresponding equilibrium carrier density n_e is given by:

$$n_e = 2.2 \times 10^{22} \exp(-0.47 \text{ eV}/kT) \qquad \text{cm}^{-3}$$

These figures coincide with the ones reported by Hoyt et al. /3/, obtained by a fitting of literature data. A deviation from the above trend is observed in Fig.2 at higher temperatures, a result which is unexpected considering that the eutectic temperature is 1097°C /4/. We point out that the value at 1100°C in Fig.2 coincides with the one after the initial 3 hours annealing at this temperature. The above equilibrium values of

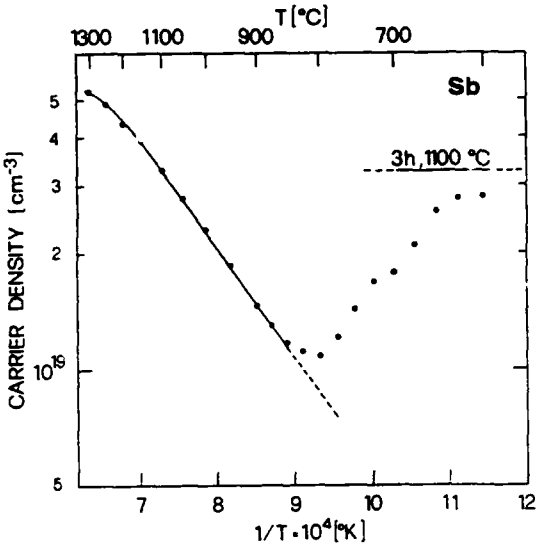


FIGURE 1

Carrier concentration vs reciprocal temperatures for Sb doped samples.

Arsenic are in very good agreement with the ones obtained by carrier profiles measurements after equilibration annealing of single crystal specimens implanted with different doses of the dopant /5/. These experiments

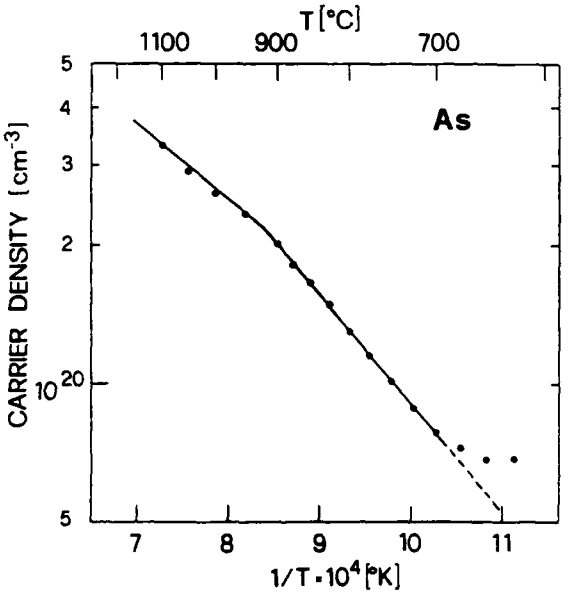


FIGURE 2

Carrier concentration vs reciprocal temperature for As doped samples.

showed that the carrier density after thermal equilibration depends only on temperature and is insensitive to excess dopant. We concluded from this results, which are supported by the occurrence of reversion and by TEM and SAXS examinations, that the equilibrium carrier density corresponds to the solubility. The same conclusion was reached also in the case of antimony by additional experiments performed on the same line, i.e. accurate carrier profile measurements after equilibration at 1100°C of single crystal specimens implanted at 160 keV with three different doses of the dopant. The results are shown in Fig.3 for an annealing time of 4 h.

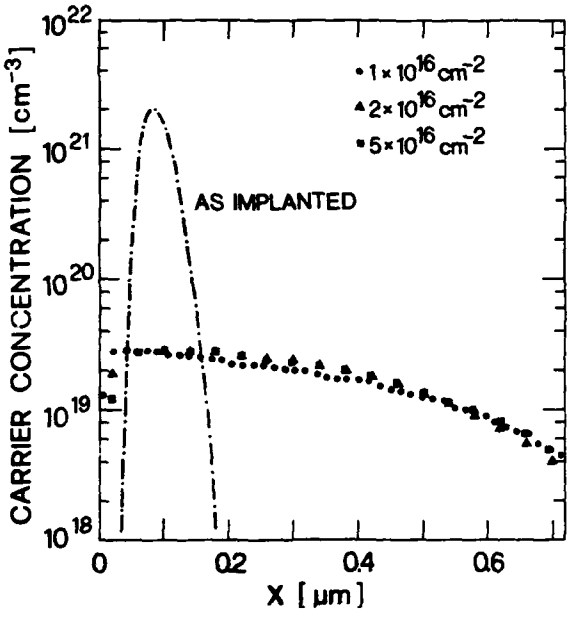


FIGURE 3

Carrier concentration profiles of Sb doped specimens implanted at different doses and annealed at 1100°C for 4 h. The as-implanted distribution for the lowest dose is also reported.

TEM observations performed on these samples evidenced the presence of a high density of Sb particles, having a size which decreases with increasing the implanted dose. According to the classical nucleation theory /7/, the density of the precipitates increased by increasing the supersaturation. In Fig.4 is reported a dark-field image of these particles, taken in a sample implanted with 2×10^{16} Sb/cm^2. They have the structure of the hexagonal antimony, as deduced from electron diffraction patterns.



FIGURE 4

Dark-field TEM micrograph, showing Sb precipitates in a 2×10^{16} cm^{-2} implanted sample, annealed at 1100°C for 4h.

High temperature data, above 1100°C, for As are not reported in Fig.2 because we verified that they were affected by the cooling rate. This phenomenon, which is attributed to additional precipitation taking place in the cooling stage, is more effective in polycristalline specimens. In fact dislocations and grain boundaries enhance the diffusion and nucleation kinetics of the dopant, a feature which, on the other side, makes polycristalline films more suitable to obtain equilibrium values in the low temperature range.

We point out that this phenomenon was not appreciably observed in antimony doped specimens, very probably due to the lower diffusivity of this dopant.

The accuracy of our solid solubility data for Sb, which correspond to the equilibrium carrier density values in Fig.1, made possible to analyze in more detail the phase equilibria for the Sb-Si system.

Experimental determinations of the liquidus curve in the phase diagram were performed by

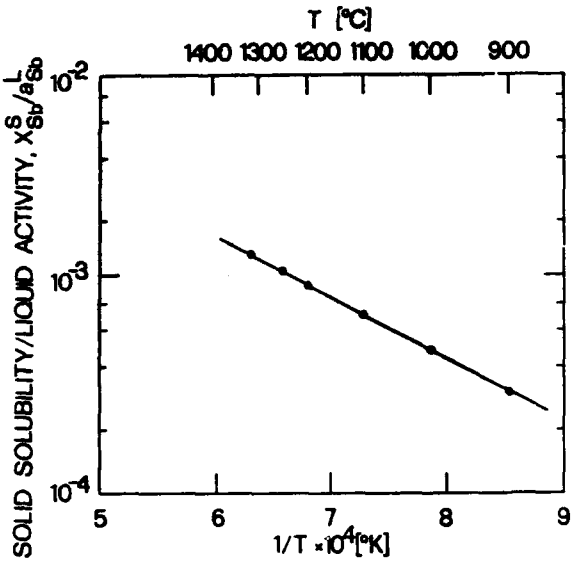


FIGURE 5

Ratio between solid solubility and liquid activity vs reciprocal temperature for antinmony in silicon.

Thurmond et al. /6/. From these data it can be deduced that a regular solution model is suitable for the liquid phase, thus allowing the calculation of the interaction parameter and hence of the activity of antimony a^L_{Sb} in the liquid /7/. The ln X^s_{Sb}/a^L_{Sb} values are reported vs reciprocal temperature in Fig.4; the precise exponential dependence which is verified in the whole temperature range is a clear confirmation of the solubility data in Fig.1. In addition this analysis provided thermodynamic data for antimony in solid solution into silicon: from the exponential dependence in Fig.5 a value of 13.4 Kcal/ mol was determined for the relative partial molar entalphy $H_{Sb}-H^o_{Sb}$; and respectively -4.7 e.u. for the relative partial molar excess entropy $(S_{Sb}-S^o_{Sb})^{xs}$.

ACKNOWLEDGEMENTS

This work was partially supported by CNR - Progetto Finalizato "Materiali e Dispositivi per l'Elettronica a Stato Solido".

REFERENCES

/1/ Nobili, D., Conf.Proc. of Satellite Symp. on "Aggregation Phenomena of Point Defects in Silicon" E.Sirtl, J.Goorissen Eds. (The Electrochem. Soc. Munich 1982) pp.189-208.

/2/ Olesinski, R.W., and Abbaschian, G.J., Bullettin of Alloy Phase Diagrams, 6 (1985) 445

/3/ Hoyt, J.L., and Gibbons, J.F., Mat.Res. Soc. Symp.Proc., "Rapid Thermal Processing", 52 (1986) 15

/4/ Olesinski, R.W., and Abbascian, G.J., Bullettin of Alloy Phase Diagrams, 6 (1985).254

/5/ Nobili, D., Carabelas, A., Celotti, G.C., Solmi, S., J. Electrochem. Soc. 130 (1983) 922

/6/ Thurmond, C.D., and Kowalchik, M., Bell Syst.Tech.J., 39 (1960) 169

/7/ Christian, J.W., "The Theory of Transformations in Metals and Alloys", Pergamon NY, Chapts. 6 and 10, 1975

DIFFUSION AND SOLUBILITY OF GOLD IMPLANTED IN SILICON

S.Coffa,L.Calcagno and S.U.Campisano

Dipartimento di Fisica,Corso Italia ,57 -Catania

G.Calleri and G.Ferla

SGS Microelettronica-Stradale Primosole - Catania

Diffusion and solubility of gold implanted in <100> p-type silicon have been investigated by Rutherford Backscattering Spectrometry and spreading resistance techniques. The gold concentration profiles are U-shaped and the concentration at the middle of the wafer thickness (C_m^s) is proportional to the square root of the diffusion time in agreement with the kick-out mechanism. The diffusion coefficient D_I^*(see text) is well described by $D_I^*=7.0*10^{-3}$ exp(-1.61/kT) cm^2/sec in the temperature range 1173-1373 K. The entropy factor associated to the ionization of the gold donor level has been determined to be 28 ± 2.

1. INTRODUCTION

The diffusion of gold in silicon has been widely investigated [1,2] because of its technological applications such as controll of the minority carrier lifetime. The electrical parameters of Au doped silicon are of great interest in silicon power devices [3]. Gold is normally diffused in silicon starting from a thin (~300 Å) layer deposited on the surface and the concentration profile is determined by the thermal process. The introduction of gold by ion implantation will result in a better control of the gold amount in the wafer,especially close to the surfaces i.e. in the electrically active region of most devices.

2.EXPERIMENTAL

P-type <100>oriented silicon, 20 Ohm-cm resistivity is used. The wafer thickness is 620 μm and double polished wafer are used to avoid gettering of gold by the rough surface. Gold implantation is performed by means of 120 KeV Au ions and the doses are in the range 10^{12}-$5*10^{15}$ atoms/cm^2. The thermal processes are carried out under nitrogen flux in the temperature range of 1173-1373 K.

Rutherford backscattering spectrometry (RBS) of 2.0 MeV He beam is used to measure the amount of gold in the near surface region (1000 Å) of the implanted wafer,the difference between the measurements before and after the thermal process resulting in the total

amount of gold diffused into the wafer. The RBS technique cannot give informations about profiles at large depths that can be instead obtained by spreading-resistance [4] technique based on the compensating effect of the two gold levels on the silicon conductivity[5].

For spreading resistance measurements the samples are mounted on a bevel block with a bevel angle of 5°44' which gives a depth resolution of 5 μm. Two tips are then leaned on the surface with a controlled pressure and the application of a small voltage makes possible to measure the spreading resistance value.

3.RESULTS AND DISCUSSION

Spreading resistance measurements (R_s) are shown in Fig.1 for 1243 K diffused samples. The resistivity (ρ) values are obtained by appropriate calibration performed by using homogeneously doped samples and the experimental data are fitted by $R_s=870\ \rho^{0.98}$. To convert resistivity into gold concentration it is necessary to solve the charge neutrality equation in the form

$$p + N_{Au}^{+} = n + N_{Au}^{-} + N_B \qquad (1)$$

where p and n are the holes and electrons concentration , N_{Au}^{+} and N_{Au}^{-} are the concentrations of the positively and negatively charged gold atoms and N_B is the concentration of boron in the substrate. The values of N_{Au}^{+} and N_{Au}^{-} depend on the entropy factors of both donor and acceptor (X_D ,X_A) gold related levels [6]. For the adopted wafer doping and gold concentration the gold acceptor level has a negligible influence on the final results and can be neglected.

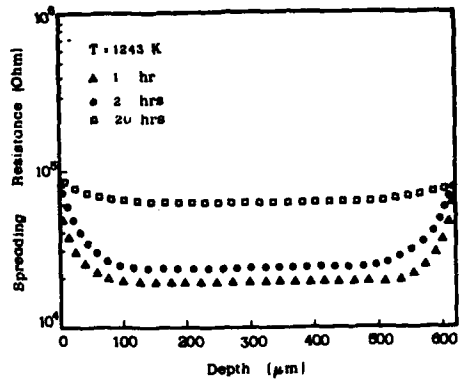


Fig.1-Resistivity profiles at 1243 K for different diffusion times.

The solution of (1) and the mobility values given in ref. 7 lead to the relation between the resistivity and the gold concentration ,which is shown in Fig.2 for different X_D values. As it appears the X_D value strongly affects the conversion resistivity - gold concentration. Because of the

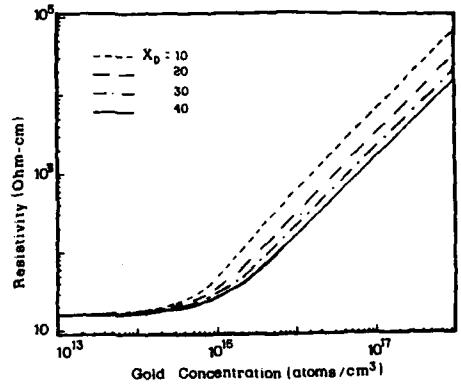


Fig.2-Silicon resistivity versus gold concentration for different X_D values.

large spread of the X_D value existing in literature [8,9,10,11] such conversion is affected by large uncertaties. Using the calculations reported in Fig.2 the resistivity profiles have been converted into concentration profiles and the total amount of gold into the wafer (area under profile) has been determined as a function of X_D as shown in fig 3. The value of the area for each thermal process is determined by RBS measurement. It is thus possible to determine the X_D value which is found to be 28±2 . Gold diffuses in silicon by migration of the fast interstitial atoms that can jump to substitutional positions. Two different mechanisms have been proposed for the interchange between interstitial and substitutional position. In the Frank-Turnbull mechanism [12] the reaction is given by

$$Au_i + V \rightleftarrows Au_s$$

where V is a vacancy. In the kick-out mechanism [13] ,

$$Au_i \rightleftarrows Au_s + I$$

where I is a silicon interstitial. The two mechanisms lead to a different trend of the gold concentration at the center of the wafer thickness (C_s^m) versus diffusion time (t). The kick-out one predicts that C_s^m increases according to

$$C_s^m = C_s^{eq} x 2/d(\pi D_I^* t)^{1/2}$$

where d is the sample thickness ,C_s^{eq} the solubility limit and D_I^* is an effective diffusion coefficient given by

$$D_I^* = D_I C_I^{eq}/C_s^{eq}$$

C_I^{eq} and D_I being the equilibrium concentration and the diffusion coefficient of silicon interstitial respectively. In Fig.4 we report C_s^m as a function of $t^{1/2}$: the agreement with the kick-out mechanism is rather good and we can estimate the effective diffusion coefficient for the investigated temperatures. An Arrhenius plot of D_I^* is reported in Fig.5 and the data are fitted by the relation

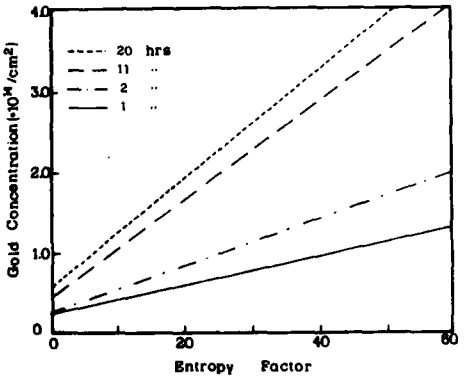


Fig.3-Total Au concentration (at./cm^2) at 1243 K as a functicn of X_D.

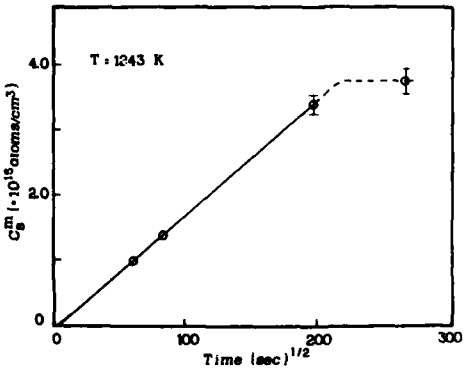


Fig.4-Gold concentration at the center of the wafer thickness (C_s^m) as a function of diffusion time for different temperatures.

$D_I^* = 7.0*10^{-3} \exp(-1.61/KT)$ cm^2/sec

shown as solid line in fig.5.

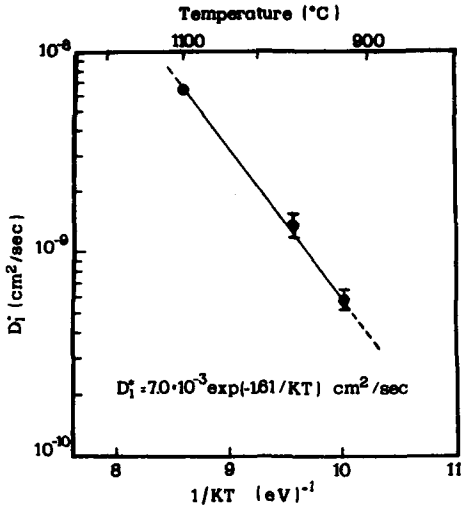


Fig.5-Arrenius plot of the effective coefficient D_I^*(see text).

4. CONCLUSIONS

By using gold implanted samples we have investigated the diffusion process of gold in silicon. Our data are in reasonable agreement with the kick-out mechanism. The activation energy for the diffusion and the entropy factor of the gold related donor level have been determined and they result 1.61 eV and 28 respectively.

ACKNOWLEDGEMENTS

The work has been supported by Progetto Finalizzato "Materiali e dispositivi per l'elettronica dello stato solido" (CNR).

REFERENCES

[1] W.M.Bullis, Solid. State El. 143 (1965) 9

[2] U.Gosele,W.Frank, A.Seeger Appl. Phys. 23 (1980) 361

[3] J.Baliga and E.Sun I.E.E.E. Trans. El. Dev. 24 (1977) 685

[4] E.E.Wang "Impurity doping processes in silicon" North Holl. Amsterdam (1981) pag. 552

[5] S.D.Brotherton, J.Bickell J. Appl. Phys. 49 (1978) 667

[6]J.A.Van Vechten and C.D.Thurmond Phys. Rev.14 (1976) 3539

[7]C.Jacoboni,C.Canali,G.Ottaviani andA.A.Quaranta Solid State El. 20 (1977) 77

[8] S.D.Brotherton,J.E.Lowther Phys. Rev.Lett. 44 (1980) 606

[9] R.Kassing, L.Cohausz, P.Van Staa, W.Mackert and H.J.Hoffman, Appl. Phys.A 34(1984)41

[10] J.A.Pals Solid. State Elec. 17 (1974) 1139

[11]N.A.Stolwijk,J.Holz and W.Frank Appl. Phys.A 39 (1986) 37

[12]F.C.Frank and D.Turnbull Phys. Rev. 104 (1956) 617

[13] U.Gosele, F.Morehead,W.Frank and A.Seeger Appl.Phys.Lett.38 (1981) 157

OPEN STENCIL MASKS FOR ION PROJECTION LITHOGRAPHY

L.-M. BUCHMANN, L. CSEPREGI, and K.P. MÖLLER

Fraunhofer-Institut für Mikrostrukturtechnik
Dillenburger Str. 53, D 1000 Berlin 33, West Germany

A processing scheme for the manufacturing of an open stencil mask has been set up by application of silicon technology and only one single X-ray lithography step for pattern generation. The mask fabrication is fully adapted to the demands of an ion projection lithography equipment by IMS. It has been proved that this mask technology permits solid structures of a complex geometry with high pattern fidelity.

1. INTRODUCTION

Ion projection lithography promises to be a successful method yielding structures in the 0.1 µm range /1/. To achieve the optimum gain from this system open stencil masks should be used which are designed to fulfill the special requirements, i.e. adapted to the ion beam divergence, the tension by thermal stress and the sputter yield by particle bombardment. A silicon membrane with a nitride top layer (Fig. 1) revealed to be most suitable for this purpose.

Considering the demands of an IMS ion projection lithography machine the fabrication of an open stencil mask will be given. Pattern generation was performed by X-ray exposition, whereas the general processing is well established in CMOS technology.

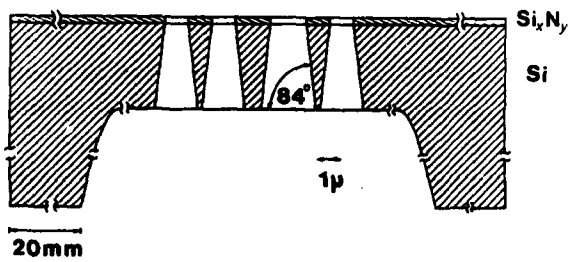


FIGURE 1

Schematic cross-section through an open stencil mask for ion projection lithography generated from a 4" silicon wafer (0.5 mm thick).

2. PROCESSING SEQUENCE

2.1 Preparation of the Membrane

The production of the thin membrane (Fig. 2) followed the same process schedule as applied for the absorber masks in X-ray lithography /2/. A P-doted 4" silicon wafer served as the substrate on which a silicon layer of 2 to 3 µm, containing boron and germanium, was deposited epitaxially $^{+}$. Subsequently, a silicon nitride layer of 0.12 µm was precipitated by a LPCVD process which is essential for pattern definition later on. Because the tensile strength between nitride and silicon layer is very pronounced additional "impurity" atoms were implanted , which lowered the stress considerably by rearranging the lattice of the nitride. When omitting this step, the membrane cracked in the concluding etching of silicon yielding debris of coiled up fragments.

The thin membrane was prepared by a two-step process. Starting with an isotropically eroding mixture of different acids (HF, HNO_3, CH_3COOH) a smooth transfer from the original backside of the wafer to the brim of the foil was formed followed by applying an ethylenediaminepyrocatechol (type S) etching agent removing silicon till the highly boron doped layer was excavated. With different covering masks windows corresponding to the geometry of the intended ion transmission area could be set.

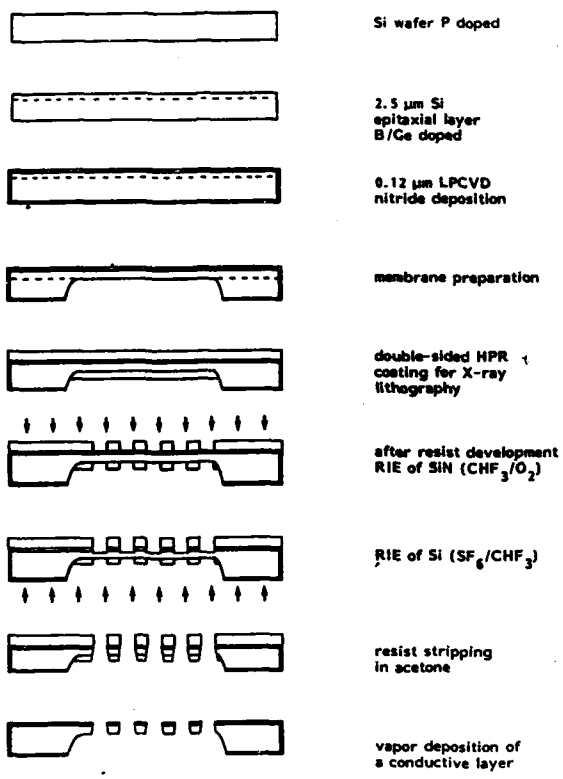


FIGURE 2

Process scheme for open stencil masks

2.2 Pattern Generation by X-Ray Lithography

The pattern was generated by exposing an HPR resist with synchrotron radiation through an absorber mask. For details concerning manufacturing of the mask and the conditions of exposure see /3/. Due to the transparency of the membrane the pattern could be transferred with coincidence into resist layers spunned on both sides of the foil by only one exposure step. The pattern fidelity of the backside test structures were routinely checked in a light-optical microscope. Fig. 3 shows a corresponding SEM micrograph.



FIGURE 3

Pattern fidelity of HPR on the backside of the membrane after exposition to synchrotron radiation and development (SEM).

2.3 Pattern Transfer - Opening of the Mask

Pattern transfer into the nitride layer of the topside of the membrane, i.e. the face the ion beam impinges on, was performed by an RIE process in a batch reactor (AME 8111). Employing a CHF_3/O_2 discharge the nitride was patterned yielding bias-free side walls without polymer depositions. Thus, the exact pattern definition could be obtained (Fig. 4).

For the purpose of mask opening a second RIE process using the fluorine containing components SF_6 and CHF_3 was determined. Etching with chlorine (BCl_3/Cl_2 or CCl_4), although generally in practice, would have required a load lock on the apparatus and, therefore, a sophisticated system to handle the fragile foils.

This process has been optimized to form the tilted sidewall corresponding to the divergence of the ion beam. The intended angle of 84° could only be obtained by sufficient sidewall passivation by species from the resist and the plasma. CHF_3 undergoes ion-molecule reactions in the plasma by which species of a high mass number (precursor to polymer formation) were originated. Therefore, the CHF_3 flow was varied in order to deposite a protecting layer, whereas the SF_6, the main source for



FIGURE 4

Pattern transfer into the nitride layer by an RIE process (CHF_3/O_2). HPR was stripped in acetone.



FIGURE 5

Sidewall protection of silicon by polymers from the resist and the CHF_3/SF_6 plasma

the highly reactive F atoms, was responsible for the amount of the etching rate. The best approach was obtained balancing these two effects with a ratio CHF_3:SF_6 of 1 (Fig. 5).

Further consideration had to be put on the thermal stress of the membrane and, especially, the resist by the etching process. Because the thermal conductivity of the delicate network is low, the transition from the wafer to the support was improved by a metal plate, to guarantee a better drain for the heat.

Due to the high selectivity of the silicon etch process with regard to the nitride an overetch could be applied to open the intended fields all over the active mask area without affecting the shape of the nitride sidewalls. Fig. 6 depicted a tapered structure obtained by etching silicon with the submitted process.

Finally, the resist was rinsed in acetone and a gold coating was deposited on the mask to prevent charging. Fig. 7 shows a 4" wafer with a 35 mm diameter membrane, which has an open area of 35%.



FIGURE 6

Tapered sidewalls of silicon test structures proposed to serve as a mask for ion projection lithography. The resist has already been dissolved in acetone.

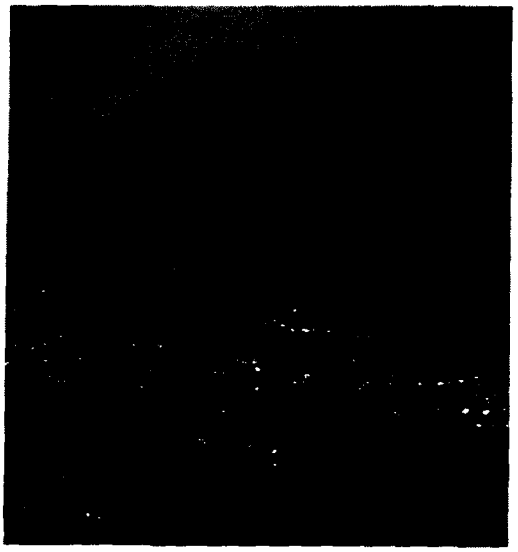


FIGURE 7

Total view on a 4" mask with an active area of 35 mm in diameter. The high transparency demonstrates the extreme thinness of the membrane.

3. CONCLUSIONS

Open stencil masks carrying test structures in the µm-range and below could be obtained by applying X-ray lithography for a self-aligned pattern generation on a membrane coated with resist on both sides and a pattern transfer splitted into two RIE processes. The exact pattern definition has been performed in a nitride layer, whereas a special profile in the supporting silicon was arranged by an etching process with compounds containing fluorine. For compensating the stress between these two layers an additional doping process of the silicon has been developed.

ACKNOWLEDGEMENTS

This work was funded by the Ministry of Research and Technology of the Federal Republic of Germany and by the Siemens AG, Munich.

NOTES

/+/ Boron in a concentration of $10^{20}cm^{-3}$ can terminate the chemical etching of silicon with an EDPS solution and thus forming the membrane. Germanium has been admixed for the purpose of internal stress compensation.

REFERENCES

/1/ Löschner, H., Ion Beam Lithography, This Conference

/2/ Hersener, J., Herzog, H.-J., Csepregi, L., Microcircuit Engineering 84, Academic Press 1985, p. 309

/3/ Heuberger, A., Microcircuit Engineering 86, North-Holland 1986, p. 3

A NEW ISOLATION PROCESS FOR VLSI DEVICES

E. Figueras*, J.L. Coppee and F. Van de Wiele

Université Catholique de Louvain, Laboratoire de Microelectronique
Place du Levant, 3
1348, Louvain-la-Neuve, Belgium

In this paper we present a new zero-bird's beak process which, with an additional photolithographic step, substitutes the thermal fully-recessed field oxide for a CVD oxide. With this process we have fabricated devices which present a very small narrow channel effect. Moreover, the use of a reference mask increases the process reproductivity and reduces the probability of the double-threshold voltage effect. The cross-section and electrical results are presented.

1.INTRODUCTION

To overcome the difficulties of LOCOS process in VLSI, several new isolation technologies [1-4], which employ a deposited oxide as a fully recessed field oxide, have been proposed. From these technologies, the BOX process [2], which uses an additionnal mask step and a double resist patterning technique prior to a planarization etch-back, is the most interesting. Nevertheless, its major drawback is that the thickness of the resist may be different for various layout designs. As a result of that, the field oxide thickness may vary in different areas [5].

In this paper, we present a new BOX process,which allows us to obtain a uniform field oxide independently of the layout design by using an additional photolithographic step and two polysilicon layers.

2.PROCESS

The starting material is <100> orientated, 14Ω.cm p-wafers. First, a pad-oxide is grown and a nitride layer is deposited. After a photolithographic step to delineate the active regions both layers are etched. The reasons to use the nitride layer will be discussed later. Then grooves are etched in the silicon substrate with a KOH solution up to a depth of 0.5 μm, using the nitride film as mask. A thermal oxidation is carried out to improve the quality of the SiO_2/Si interface. After the channel-stop implantation, a SiO_2 layer, with an equal thickness (or little more) to that of the Si groove depth, is deposited by CVD[1] (Figure 1.a). A polysilicon layer is deposited and etched by plasma with the help of an additional resist mask which protects the polysilicon in the field areas (Figure 1.b). The grooves between the polysilicon 1 and the oxide walls are refilled with a second polysilicon layer (Figure 1.c). Figure 1.d shows the stucture after polysilicon 2 etch-back and SiO_2 etching. During the last step, the nitride, the polysilicon and the pad-oxide layers are removed to obtain the desirable structure.

Figure 2 illustrates the utility of the nitride mask. This layer is used as a reference mask. After it has been removed we are sure that all the active areas in the wafer are recovered with the same oxide thickness, independently of the misuniformity during polysilicon and oxide etching steps. The better the uniformity is, the thinner the silicon nitride layer can be. The choice of the reference mask material is limited by three drawbacks: it may not be oxide, it must resist KOH solution, it must be stable at high temperature. Therefore we have chosen silicon nitride.

Moreover the reference mask must assure a

*E. Figueras is financially supported by the National Microelectonic Center (Education and Research Ministry, Spain)

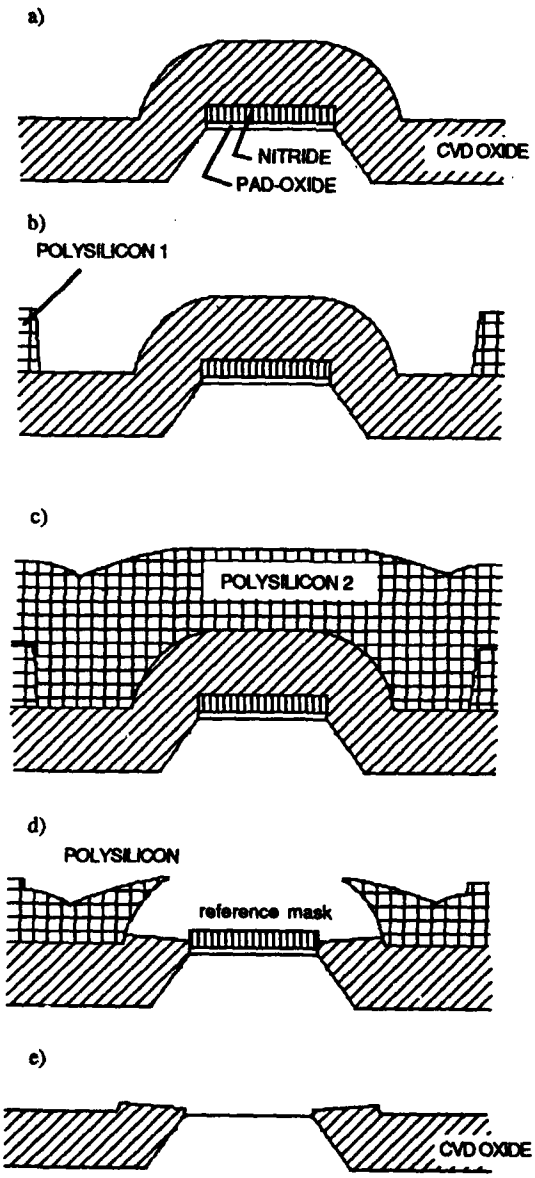


FIGURE 1

Fabrication process steps.

small positive transition step between the active and the field areas which will reduce the double threshold effect[5].

3.RESULTS

The SEM cross-sections, Figures 3 and 4, show the structures obtained following our process. It

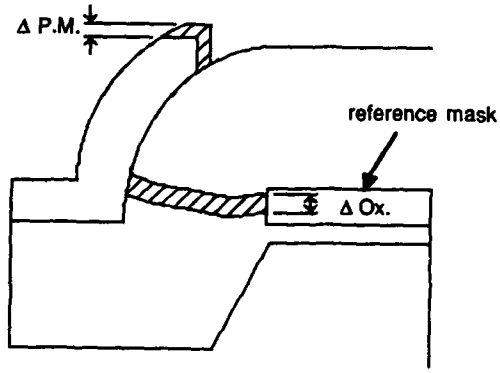


FIGURE 2

The reference mask serves to correct the misuniformity produced during polysilicon mask (ΔP.M.) and oxide (ΔOx.) etching steps.

can be seen that there's no bird's beak formation and that the field oxide thickness is uniform (except at the edge).

To present the electrical results of BOX technology, we will compare them with the results of SILO technology, wich furnish better characteristics than LOCOS[6]. Both technologies were carried out with the same test circuit and with natural transistors wich present a more important narrow-channel effect than enhanced ones due to the lower doping level of the channel.

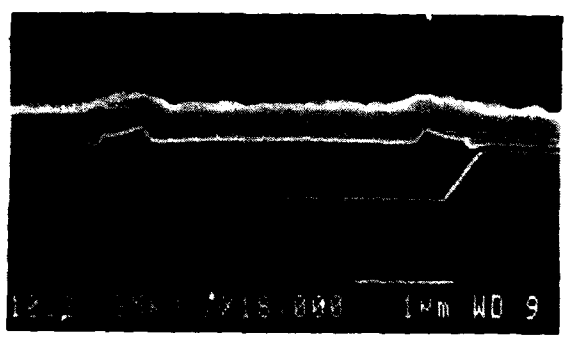


FIGURE 3

Field oxide area. The top of the structures is covered with 0.3 μm polysilicon to evidence the oxide profile.

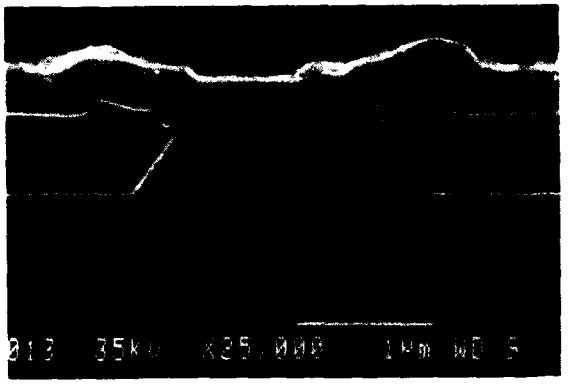


FIGURE 4

Active area with W= 1.3 µm.

TABLE I		
PROCESS	DOSE (cm^{-2})	THRESHOLD (V)
SILO 1	2 E 13	12.4
SILO 2	1 E 13	6.5
BOX	1.6 E 12	9.7

Table I shows the field implantation dose and the field oxide threshold voltage for each structure. The threshold voltage is defined at I_D=1E-6 A for a parasitic field transistor with W/L= 100 µm/20 µm.

From Table I, we can deduce, that for the same field threshold voltage, the dose used in BOX is ten times lesser.Thereby the narrow-channel effect is nearly suppressed, as shown in Figure 5.

The subthreshold curves (Figure 6) are plotted for different drain polarisation (V_D= 5.0v and 0.5v) and clearly show the absence of double-threshold effect[7], which can appear in the subthreshold characteristics of devices isolated with a BOX field oxide[1].

4.CONCLUSIONS

We have presented a new technology viewing to obtain a field recessed CVD oxide, ensuring the oxide uniformity independently of the layout.Moreover the use of a reference mask

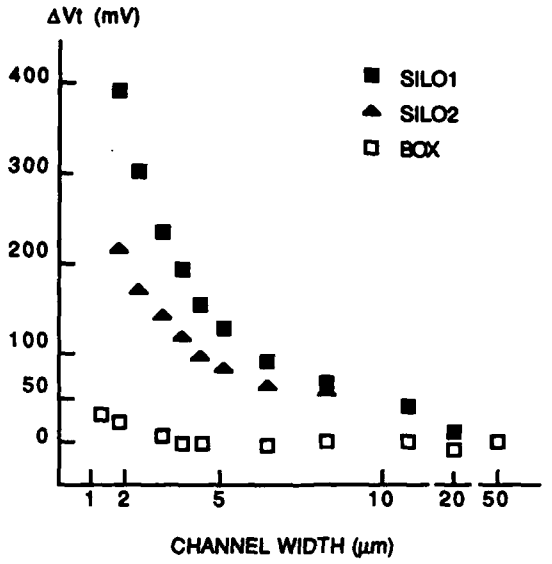


FIGURE 5

Narrow-channel effect. The respective field oxide threshold are, BOX: 9.7v, SILO1: 12.4v, SILO2: 6.5v.

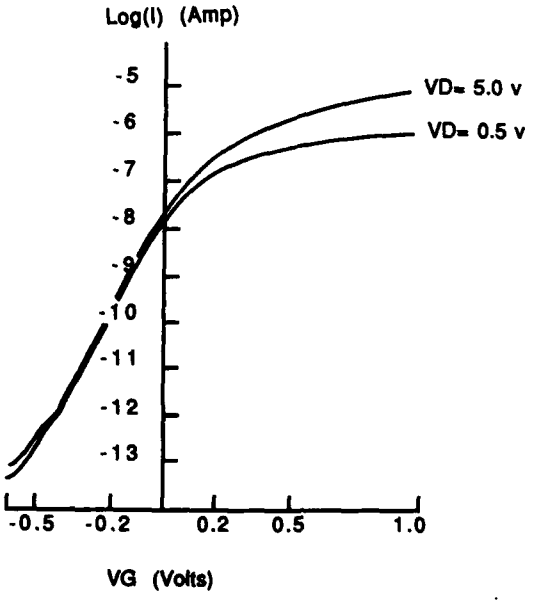


FIGURE 6

Subthreshold characteristics of a transistor with W/L= 1.9 µm/20 µm. No double-threshold effect is shown. VS= 0v, VB= 0v.

eliminates double-threshold effect because it prevents silicon steps in the active area edge. All that prove the advantages of using CVD oxide as field oxide for VLSI circuits.

ACKNOWLEDGEMENTS

We will express our sincere gratitude to the technical staff of our laboratory for their effort to carry out this process, and Dr. F. Serra-Mestres of the C.N.M.(Spain) for his firm support and encouragement.

NOTES

(1)We use a PYROX machine to deposite CVD oxide.

REFERENCES

[1] Kurosawa,K., Shibata,T., Iizuha,H., A new bird's beak free field isolation technology for VLSI devices, IEEE IEDM Tech. Dig.,paper 16.4,1981.

[2] Shibata,T., Nakayama,R., Kurosawa,K., Onga, S., Konaka,M., Iizuka,H., A simplified BOX (Buried OXide) isolation technology for Megabit dynamic memories, IEEE IEDM Tech. Dig., paper 2.3,1983.

[3] Chen,J.Y., Henderson,R.C., Hall,J.T., Yee, E.W., A fully reassed field isolation technology using photo-CVD oxide, IEEE IEDM Tech. Dig. ,paper 9.5, 1982.

[4] Yachi,T., Serikawa,T., J. Electrochem. Soc., vol 123, No 11 (1985), 2775.

[5] Oldham,W.G., Shachman-Diamand,Y., Pai,P. L., Young,K., Sutardja,P., MOS isolation technology, Proceedings of the 14th European Solid State Device Research Conference (ESSDERC), 1984, p.53.

[6] Iizuka,T., Chiu,K.Y., Moll,J.L., Double threshold MOSFET's in bird's beak free structures, IEEE IEDM Tech. Dig., paper 16.3, 1981

[7] Coppee,J.L., SILO Isolation Technique: A Study of Active and Parasitic Characteristics with Semi-recessed and Fully-recessed Field Oxides, in print.

OXIDATION INDUCED LOCAL CHANNEL DOPANT ACCUMULATION

C. Mazuré and M. Orlowski

Siemens AG, Corporate Research and Development
Otto-Hahn-Ring 6, D-8000 München 83, West-Germany

The influence of the reoxidation after gate patterning on NMOS transistor characteristics is investigated. As the channel length L_G is reduced the devices exhibit a reoxidation related V_T increase. This V_T variation is explained by a non-uniform oxidation enhanced diffusion of the channel dopant along the channel.

INTRODUCTION

In order to maintain long channel behaviour at submicron channel lengths ($L_G \leq 1.0$ µm) high doping concentrations in the channel region are required. We consider the case where this is achieved by a double channel implant. For punchthrough reduction a high dose deep implant is used. Consequently the channel profile exhibits a maximum in the bulk.

After the gate patterning etch a reoxidation step may follow prior to or after the LDD implantation. In this work we focus on the NMOS transistor characteristics due to the influence of this particular oxidation.

DEVICE FABRICATION

n^+-poly-Si-gate n-channel transistors with conventional LDD technique were used. The gate oxide thickness was 20 nm. After gate patterning a dry reoxidation at 900°C was performed. The following cases were investigated: a) a 40 min reoxidation prior to phosphorus-LDD implant and b) a 8 min reoxidation before and after the LDD implant. In both cases the total processing time at 900°C was the same. The junction depth was approximately 0.35 µm.

EXPERIMENTAL RESULTS AND DISCUSSION

It is found that the threshold voltage V_T increases as L_G is reduced in the case of reoxidation. The V_T increase for different channel implants is shown in Fig.1. Note that there is a V_T increase already for $L_G \leq 3$ µm. It is a long range effect. In Fig. 2 the effect of the reoxidation time t_{REOX} on V_T is shown. The channel implants are the same in both cases. The difference dV_T refers to the maximum V_T value minus the long channel (L_G = 10 µm) V_T value.

In principle a V_T increase could be related to the bird's beak at the gate

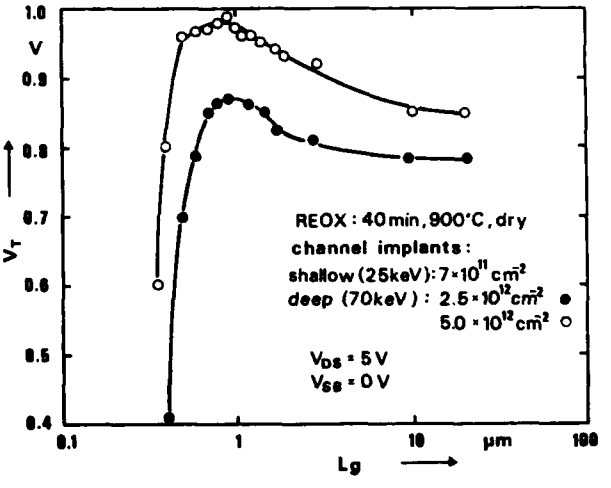


FIGURE 1:
Threshold voltage vs channel length

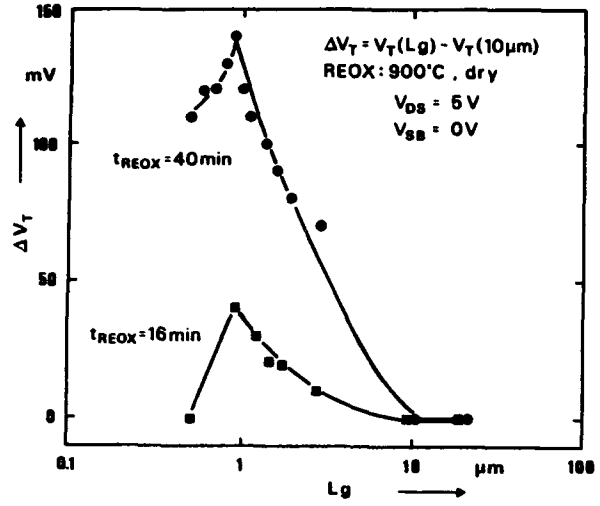


FIGURE 2:
Threshold voltage variation vs channel length.

edge due to the reoxidation. An increase of t_{REOX} could induce a further V_T increase. But it does not fit to the long range character of the effect we observe. Furthermore the bird's beak lies within the n⁻-region. Samples made without channel implant in highly doped substrates showed no V_T increase for channels as short as L_G = 1 μm for t_{REOX} = 40 min before short channel effects took over.

We explain the V_T enhancement by a non-uniform dopant redistribution along the channel region near the surface due to the reoxidation step. During gate reoxidation a local oxidation enhanced diffusion (OED) /1/ occurs. The channel dopant diffusion is not only enhanced under the oxidized region but also in the adjacent unoxidized channel area. The channel profile has its maximum in the bulk as shown in Fig.3. Therefore the lateral injection of the institials leads to a local enhancement of the boron diffusion normal to the surface. We expect the lateral increase of the surface concentration C_s along the channel to vary according to the lateral interstitial decay length l_x. A schematic diagram is shown in Fig.4. It follows that even for very long channels V_T will be slightly enhanced compared with the laterally uniform channel profile case. A similar V_T behavior has been observed for DI-LDD transistors due to the deliberate boron halo implantation, however for much shorter channels, $L_G \leq 0.5$ μm /2/.

As L_G is reduced below $2 \cdot l_x$ the regions of enhanced C_s are expected to merge leading to a further V_T increase. The onset of the V_T increase determines the lower bound for l_x. From Fig. 1 and 2 we have $2 \cdot l_x > 3$ μm in agreement with the values reported in the literature for l_x /1,3/.

In Fig.3 the simulated 1-D vertical doping profiles are shown for the case where the oxidation is treated as an inert anneal (solid line), and when OED

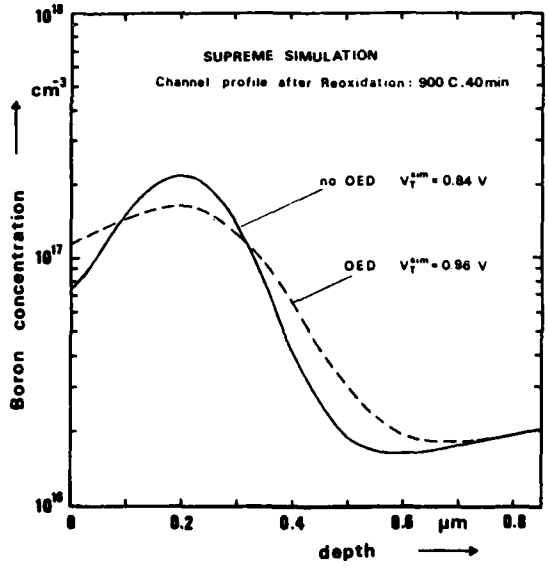


FIGURE 3:
Simulated vertical profiles with and without OED.

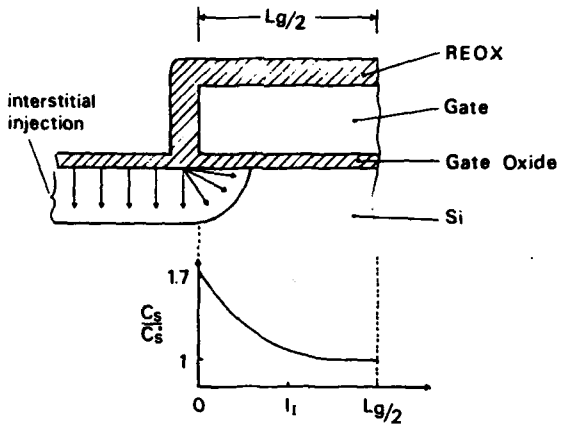


FIGURE 4:
Schematic diagram of lateral C_s.

is included (dashed line). The latter profile represents thus the channel profile near the gate edges. The solid line profile, on the other hand, corresponds to the vertical profile halfway between source and drain for $L_G > 2*l_I$. The experimental data for dV_T and the simple model predictions are in satisfactory agreement (Fig.5 and 6).

Since the concentration gradient $\delta C_s/\delta x$ is the driving force for the doping redistribution, the local increase or decrease of C_s depends on the magnitude and sign of the gradient of the channel profile at the surface. In our case (Fig.3) the positive gradient will induce a lateral increase of C_s which acts as an additional surface punchthrough stop.

Therefore by controlling the concentration gradient it should be possible to control dV_T. Fig. 5 shows the dV_T dependence on the implant energy E_B^d of the deep implant for a constant implant dose. With increasing E_B^d the concentration gradient diminishes, thus reducing dV_T . In Fig. 6 the calculated dV_T as function of E_B^d and of the concentration gradient is shown. Notice the good qualitative agreement with the experimental data from Fig. 5. There is a tight correlation between dV_T, t_{REOX} and the gradient of the channel profile at the surface.

As a positive gradient helps improve the short channel behaviour, the opposite happens for a negative gradient. In this case the short channel

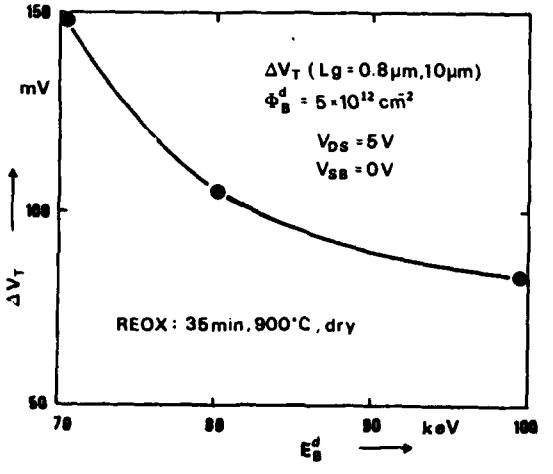


FIGURE 5:
Threshold voltage variation vs channel implant energy.

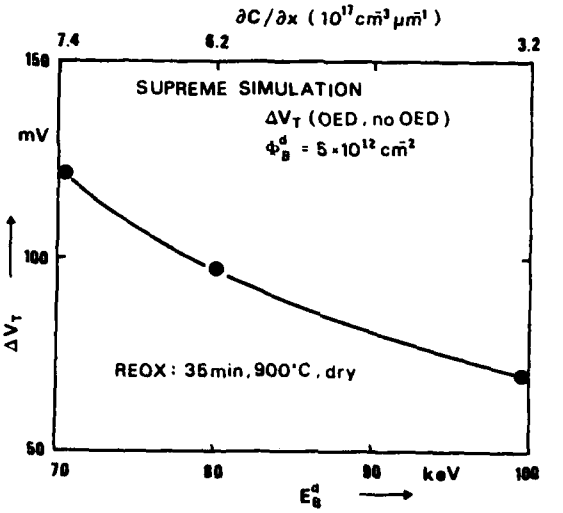


FIGURE 6:
Simulated threshold voltage variation vs channel implant energy and concentration gradient.

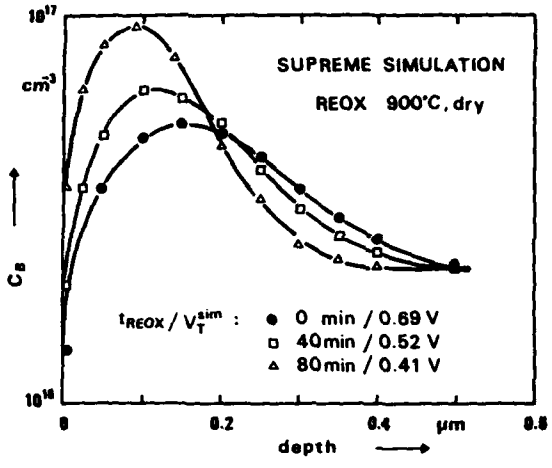


FIGURE 7:
Simulated OED efects for a shallow channel implant.

effects are enhanced. Fig. 7 shows an example of the channel implant with its concentration maximum at the surface. In this case the oxidation step would deteriorate the $V_T(L_G)$ behaviour.

SUMMARY

We have shown that a reoxidation step after gate patterning induces a lateral non-uniform redistribution of the channel dopant. This redistribution is due to the local OED effect which affects strongly the short channel behavior of NMOS transistors. We have demonstrated that there is a tight relation between the process parameters and the short channel behavior which can be technologically controlled.

REFERENCES

/1/ Taniguchi, K. , Antoniadis, D.A. and Matsushita, Y., Appl. Phys. Lett. 42 (1983), 799

/2/ Codella, C.F. and Ogura, S. IEDM Tech. Dig. (1985), 230.

/3/ Lin, A.M., Dutton, R.W. and Antoniadis,D.A., Appl.Phys. Lett. 35 (1979), 799.

REDISTRIBUTION OF FLUORINE FROM BF_2^+ IMPLANTS IN MOS STRUCTURES

Harry J. WHITLOW, Juhani KEINONEN*, Carina ZARING, and C. Sture PETERSSON

The Royal Institute of Technology, Department of Solid State Electronics, Box 70033, S-10044 Stockholm, Sweden.

*Accelerator Laboratory, University of Helsinki, Hämeentie 100, SF-00550 Helsinki 55, Finland.

The thermal redistribution of fluorine from 111 keV BF_2^+ molecular ion implants in (100)Si and 56 nm thick thermally grown SiO_2 films on Si substrates has been studied using the Nuclear Resonance Broadening (NRB) technique. Vacuum annealing at temperatures up to at least 700°C lead to no loss or redistribution of fluorine from either the pure Si or oxidised Si samples. Fluorine diffusion took place for temperatures of 900° C and above. At these temperatures fluorine was lost from the Si samples but agglomerated, most probably at the SiO_2/Si interface in the oxidised samples.

1. INTRODUCTION

In recent years BF_2^+ molecular ion implantation has been widely employed in Metal Oxide Semiconductor (MOS) technology for the production of shallow B-doped layers. These shallow doped layers are important not only for the source and drain regions but also BF_2^+ implants in SiO_2/Si structures are used for threshold voltage control and the fabrication of high-capacitance cells for memory applications[1].

The fluorine that is co-implanted with the boron dopant constitutes a potentially troublesome impurity since fluorine segregation to the SiO_2/Si interface is associated with degradation of the breakdown characteristics of MOS structures[1-3]. It is generally assumed that fluorine is lost even from capped samples during thermal annealing at 900°C and above. However no unambiguous study has to our knowledge been reported but a recent study on capped samples[4] suggests that segregation of fluorine to the SiO_2/Si interface might take place.

Recently Transmission Electron Microscopy (TEM) studies[5,6] show that gas bubbles are formed during annealing at elevated temperatures. The presence of such TEM visible gas bubbles implies that profiling techniques that rely on sputter-erosion will give erroneous profiles and early work should therefore be critically examined. Moreover in MOS structures the fluorine distribution may differ significantly from the case of implantation into pure-Si since chemical driving forces and radiation enhanced mobility[7] may significantly modify the redistribution of fluorine (and boron).

2. EXPERIMENTAL

2×10^{15} 111 keV $^{11}B^{19}F_2^+$ ions-cm^{-2} were implanted into (100)Si wafers that were cleaned by dipping in a HF 1:10 solution, as well as (100)Si wafers that

were thermally grown to form a thin oxide layer,(to mimic a gate oxide). The wafers were subsequently divided and vacuum annealed at temperatures between 550°C and 1000°C for 30 min.

The thickness of the oxide layers on the HF-dipped and oxidised Si samples were determined using mass and energy dispersive recoil spectrometry[8] to be 2.5 and 56 nm respectively assuming bulk densities and stoichiometry.

Fluorine was quantitatively profiled in the structures using the Nuclear Resonance Broadening (NRB) technique. The 340 keV proton energy resonance in the $^{19}F(p,\alpha\gamma)^{16}O$ reaction was used, which is well suited for depth profiling fluorine with the high sensitivity required for impurity profiling in microelectronic structures[9-12]. Protons from the University of Helsinki 2.5 MV Van de Graaff accelerator were used to excite the reaction, and the emitted γ-rays were detected in a scintillation counter system[11,12].

3. RESULTS AND DISCUSSION

No change in the fluorine depth distributions were observed in either the HF-dipped or oxidised samples as a consequence of thermal annealing in vacuum up to 700°C for 30 min, which is higher than the temperature required to achieve epitaxial regrowth and a reasonable degree of electrical activation. The co-implanted fluorine is however clearly mobile in both sample structures during annealing at 900°C, especially in the region where most of the implantation induced damage resides. In the more extreme case of the 1000°C anneal where the fluorine is very mobile, there is a marked difference in the fluorine depth profiles from the Si and the SiO_2/Si samples. About 90 at.% of the co-implanted fluorine in the Si sample was lost through the surface. (The low solid solubility of F in crystalline Si rules out diffusion to greater depths.) No significant fluorine loss is observed from the BF_2^+ implanted SiO_2/Si structures at this temperature but the fluorine coalesces to a thin layer. Most probably the fluorine has segregated to the SiO_2/Si interface, however the proton energy determination was insufficient to allow us to completely rule out segregation to the SiO_2 surface.

REFERENCES

[1] K.W. Teng, H.H. Tseng, B.Y. Nguyen and F.T. Liou, J. Electrochem. Soc. extended abstract 87-1(1987)328.

[2] J. Belsen and I.H. Wilson, Nucl. Instrum. and Meth. B1(1984)374.

[3] M. Fukumoto and T. Ohzone, Appl. Phys. Lett. 50(1987)894.

[4] M.C. Ridgeway, P.T. Scanlon and J.L. Whitton, J. Appl. Phys. (In press)

[5] C.W. Nieh and L.J. Chen, J. Appl. Phys. 60(1986)3114.

[6] C.W. Nieh and L.J. Chen,Appl. Phys Lett. 48(1986)1528.

[7] H.H. Andersen, in;Ion Implantation and Beam Processing. (eds.) J.S. Williams and J.M. Poate, (Academic Press, Australia, 1984)p.127.

[8] H.J. Whitlow, G. Possnert and C.S. Petersson, Nucl.Instr. and Meth. Section B (in press).

[9] A. Anttila and J. Keinonen, Int. J. of Appl. Rad. Isotop. 24(1973)293.

[10] A. Anttila, S. Brandenburg, J.Keiononen and M. Bister, Nucl. Phys. A334(1980)205.

[11] D. Dieumegard, B. Maurel and G.Amsel, Nucl. Instrum. and Meth. 168(1980)93.

[12] H.J. Whitlow, Th. Eriksson, M. Ostling , C.S. Petersson, J. Keinonen and A. Anttila, Appl. Phys. Lett. (In press).

DOPANT REDISTRIBUTION FROM ION IMPLANTED WSi_2 ON POLY-Si

S Nygren [a], D Levy [b], G Goltz and J Torres

C.N.E.T./C.N.S., Chemin du Vieux Chêne, BP 98, F-38243 Meylan, France

The possibility of doping the polycrystalline silicon in a polycide configuration by implantation into the silicide and subsequent thermal diffusion has been evaluated. In a first step it is demonstrated that a phosphorus implantation to a dose of 5E15 cm^{-2} into WSi_2 can produce flatband voltages similar to those obtained by conventional diffusion doping into the polysilicon. This is true even if 2E15 cm^{-2} boron is subsequently implanted. Secondly arsenic and boron implantations were tried with the aim to produce n- and p-doped gates respectively. Work function shifts between the two species approaching 1 V indicate that typical source/drain implantations may be used for this purpose.

1. INTRODUCTION

One consequence of the continuing shrinkage of VLSI geometries is that heavily doped polycrystalline silicon can no longer provide the low resistive interconnects required to realize the potential performance improvements offered by reduced channel lengths. This condition has been foreseen for some time, and double layers of poly-Si and a refractory metal silicide are now coming into use. In this configuration the silicide acts as a shunt, reducing the overall sheet resistivity without influencing the well behaved and well characterized poly-Si/SiO_2 interface. As a result of this approach the need for very high doping levels in the polysilicon layer, to lower its resistivity, can be relaxed, and implantation doping may become applicable.

Another consequence of the reduction of lateral dimensions is that n^+ polysilicon may have to be replaced by a material with a higher work function as the gate material for p-channel transistors. The reason is that it is becoming increasingly more difficult to at the same time adjust the threshold voltage and avoid short channel behaviour in these devices. For CMOS technology a possible solution to this growing concern may be the use of p^+ polysilicon for the p-channel transistors, whereas n-channel gates remain n^+ polysilicon.

For process simplicity an ideal condition would be if such an n^+/p^+ design could be realized by the source/drain implants. In particular the corresponding doses have to be adequate to set the electrical characteristics of the gates to proper and reproducible values. For silicide/polysilicon structures a desired process sequence would include implantation after silicide deposition. The dopants should preferably be driven into the underlying silicon and activated by a subsequent, already existing, annealing step. Promising results in this field have been reported for $TaSi_2$ [1, 2] and WSi_2 for shallow junction formation [3]. This paper presents results

a) Permanent address: Royal institute of Technology, Solid State Electronics, S-10044 Stockholm, Sweden
b) Permanent address: Bull, Rue Jean Jaurès, F-78340 Les Clayes-Sous-Bois, France

for a polycide configuration having tungsten silicide as the top layer. The possibility of simply replacing conventional $POCl_3$ diffusion doping by a phosphorus implantation into silicide was also tested, and the influence of an additional boron implantation was studied.

2. EXPERIMENTAL

Wafers used for the phosphorus implantation investigation were (100) oriented, p-type silicon. The layered structures were formed by thermal oxide growth, LPCVD polysilicon deposition and cosputtering of the silicide film. Typical thicknesses were 25, 230 and 150 nm respectively. Prior to silicide deposition reference wafers were doped by $POCl_3$-diffusion and wafers to be implanted were annealed at 850 °C for 30 min in O_2. The latter had phosohorus implanted into the silicide at 100 keV and to doses of 5E15 and 1E16 cm^{-2}. The as deposited amorphous silicide had a stoichiometry close to $WSi_{2.3}$ as measured by RBS. A transformation into a polycrystalline film was achieved by rapid thermal annealing at 900 °C for 20 s in N_2 and 2E15 cm^{-2} boron was implanted at 30 keV. This step simulated the compensating implantation to which p-channel gates are subjected in a self aligned CMOS process. An introductory examination revealed that identical electrical results were obtained wether or not the crystallization anneal was performed. 1·1 mm^2 capacitors were patterned and etched and the polycide structures were oxidized resulting in a remaining poly-Si thickness of 150 nm. Following a high temperature anneal at 1100 °C for 20 s and removal of the capping oxide the back surface was prepared by etching, aluminum deposition and annealing.

A slightly different procedure was employed for the wafers produced for work function determination. Gate oxides were grown to various thicknesses and etched in buffered HF to form bevels ranging from 15 to 60 nm. Polysilicon annealing, silicide crystallization and polycide oxidation were excluded. The polysilicon thickness was therefore 150 nm as deposited. Silicide deposition was followed by arsenic or boron implantation at 180 and 30 keV respectively and to doses in the range from 1E15 to 1E16 cm^{-2}. The chosen energies correspond to typical values for shallow junction formation. A 60 nm thick encapsulating oxide was deposited by PECVD to prevent loss off dopants during the subsequent 1100 °C, 20 s anneal. Capacitor formation concluded the fabrication.

Samples for SIMS and RBS analysis were cut from unpatterned but identically processed wafers.

3 RESULTS AND DISCUSSION

Arsenic redistribution was recorded by backscattering spectrometry (RBS). The situation after a 10 s anneal at 1100 °C is shown in Figure 1. It is clear that most of the impurity atoms have diffused out from the silicide into the polysilicon and to the interface between the two layers. The result after 20 s at the same temperature was almost identical. This behaviour has previously been observed by Jahnel et al [4]. Similar characterizations by means of SIMS analysis are in progress for the boron and phosphorus distributions. Pan et al [5] have presented evidence for a high affinity for boron in WSi_2.

Automatic mapping of flatband voltages was performed to compare the differently phosphorus doped polycide capacitors. Capacitance-voltage characteristics were measured for 65 capacitors on each wafer and three wafers were analysed for each doping procedure and dose. Whether the impurity had been introduced by

conventional diffusion into the poly-Si or by implantation into the silicide flatband voltages were similar. This holds also for the low dose implantation (5E15 cm^{-2}). In Figure 2 the range of VFB values is indicated for the different conditions. The appearently greater dispersion displayed by the diffused capacitors may to some extent be attributed to the larger number of good capacitors measured in this case. The cause for the poor yield exhibited by the implanted wafers was not unambiguously identified. An equivalent concentration for the $POCl_3$ samples was measured by SIMS analysis and comparison with a reference sample of known concentration.

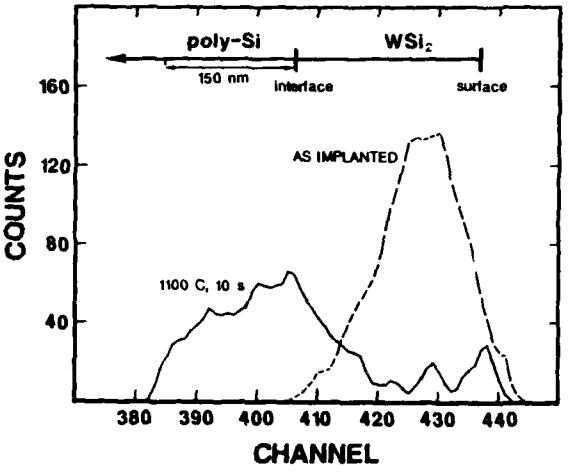


FIGURE 1. *RBS spectra showing the arsenic distribution as implanted into the silicide and after annealing at 1100 °C for 10 s.*

Wafers with bevelled gate oxides were subjected to the C(V) mapping and oxide thickness as well as flatband voltage were recorded for each capacitor. If oxide charges are limited to fixed oxide charges at the Si/SiO_2 interface, flatband voltage depends linearly on oxide thickness. The difference in work function is then given by the intercept at $t_{ox}=0$ and the slope represents the fixed charge concentration. Figure 3 is such a plot for a wafer implanted with arsenic to a dose of 1E15 cm^{-2}. The excellent alignement of the points is a good indication of the validity of this method. However there is a shift in the absolute value of the work functions and a study of its origin is under way. Nevertheless some interesting information may be deduced from an examination of the relative shifts for the different conditions. The lowest ϕ_{MS} value, representing the highest arsenic dose, was taken as a reference point and set to zero in Figure 5. For arsenic doping all values are less than 0.1 eV from this value. The recording for the 4E15 cm^{-2} dose can be explained by a slight falloff from the equilibrium depletion capacitance resulting in higher calculated flatband voltages for these capacitors. Boron implanted wafers display an increase in ϕ_{MS} with increasing implantation dose. The shifts range from 0.94 to 1.07 eV relative to the reference.

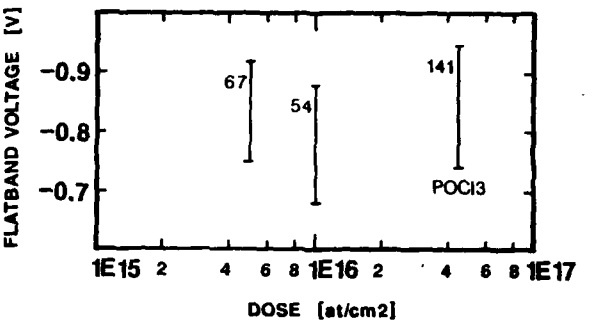


FIGURE 2. *Flatband voltages for three different phosphorus concentrations. The lower doses were implanted into the silicide and subsequently diffused by annealing at 1100 °C for 20s. The highest dose was directly diffused into polysilicon from a POCl3 source. For each condition the number of good capacitors is indicated.*

4 CONCLUSIONS

The possibility of replacing conventional phosphorus diffusion doping by implantation into the silicide and subsequent drive in annealing in a tungsten based polycide stack was examined. A comparison of flatband voltages indicate that a phosphorus dose as low as 5E15 cm^{-2} partially compensated by 2E15 cm^{-2} boron

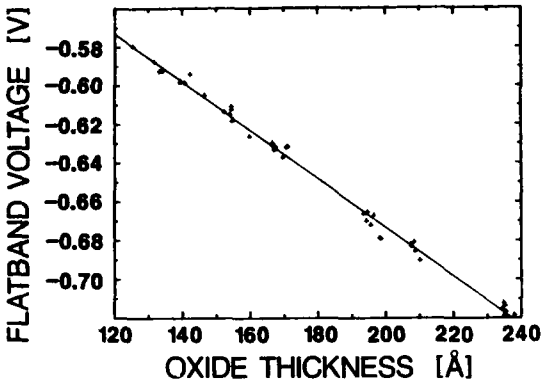


FIGURE 3. *Flatband voltage vs. oxide thickness for a wafer implanted with arsenic to a dose of 1E15 cm-2 and annealed at 1100 °C for 20 s.*

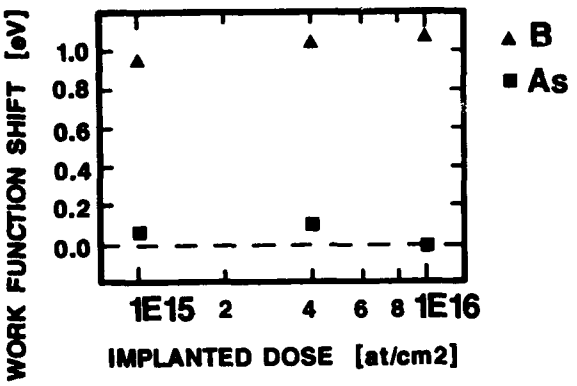


FIGURE 4. *Work function shifts for various arsenic and boron doses implanted into silicide. The highest arsenic dose was arbitrarily taken as a reference.*

is sufficient to produce electrical characteristics similar to those obtained for $POCl_3$ doping of the polysilicon layer. A similar investigation of arsenic and boron implanted structures revealed that work function differences of about 1 eV between the two cases could easily be obtained. Reproducibility was excellent over the wafers. This doping procedure thus seems to be a possible means of realizing the n^+/p^+ gate CMOS without increasing the doses beyond what is typically being used for source and drain implantations.

REFERENCES

[1] S Vaidya, T F Retajczyk and R V Knoell, *J Vac Sci Technol*, **B3**, 846 (1985)

[2] H Gierish, F Neppl, E Frenzel, P Eichinger and K Hieber, *J Vac Sci Technol*, **B5**, 508 (1987)

[3] F C Shone, K C Saraswat and J D Plummer, in *Int El Dev Meet Tech Dig*, p 407 (1985)

[4] F Jahnel, J Biersack, B L Crowder, F M d'Heurle, D Fink, R D Isaac, C J Lucchese and C S Petersson, *J Appl Phys*, **53**, 7372 (1982)

[5] P Pan, N Hsieh, H J Geipel and G J Slusser, *J Appl Phys*, **53**, 3059 (1982)

BIPOLAR HETEROJUNCTION TRANSISTORS: OUT OF MODELS INTO REALITY

A. J. Holden

Plessey Research Caswell Limited
Allen Clark Research Centre, Caswell, Towcester, Northants, NN12 8EQ, U.K.

ABSTRACT

The modelled and measured capabilities of the Bipolar Heterojunction Transistor are reviewed. The device is shown to have potential in high speed digital ICs and in microwave power amplification. It has demonstrated a world beating record for high speed divider circuits but is being hotly pursued by the Silicon Bipolar Junction Transistor. On paper it has promised much and it is now becoming a reality.

1. INTRODUCTION AND OVERVIEW

The Bipolar Heterojunction Transistor (BHJT) offers a unique combination of advantages. It draws on existing bipolar logic integrated circuit design and performance, the speed of GaAs and the gain enhancement of a wide band-gap emitter. In this paper we review the origins and basis of the device and assess its progress.

1.1 Why Bipolar?

For logic integrated circuits Si bipolar ECL offers the fastest switching available (albeit at a cost of high power consumption), compared with rival structures such as CMOS. ECL divide by four circuits have been reported, toggling at 9.1GHz [1]. Bipolar transistors also display superior threshold voltage uniformity compared to field effect devices. For microwave power amplification the bipolar transistor provides the highest output power at the highest frequencies (70W (pulsed) at 3.5GHz [2] and 2W at 9GHz [3]) compared with its generic rivals such as MOSFET amplifiers and (at least at lower frequencies) MESFETs.

1.2 Why GaAs Bipolar?

The ultimate speed of a bipolar transistor is determined by the transit time of carriers through the base and collector and so enhanced mobility is an advantage in GaAs compared with Si. The availability of a semi-insulating substrate reduces stray capacitance in circuits and reduces transmission delays introduced by 'lossy' Si substrates. A very important advantage of GaAs is the possibility of including optical functions such as optical interconnects, lasers and detectors on the same chip as the transistor. Optical interconnects offer speed and complexity advantages over standard electrical connections. GaAs also has the peripheral advantage of radiation hardness. The disadvantages include reduced carrier lifetime in GaAs, and poor thermal conductivity.

1.3 Why Heterojunction?

Homojunction bipolar transistors cannot combine high current gain with low base resistance. In order to enhance the gain the ratio of emitter to base doping must be high. Upper limits on doping depend on the technology and also loss of mobility and structure. Conversely the base doping may not be too low because of the effect on intrinsic base resistance. In Si BJTs emitter dopings in excess of $10^{20}cm^{-3}$ are feasible but in GaAs only mid $10^{18}cm^{-3}$

is more commonly achieved. To add a further dimension to the design, AlGaAs can be grown lattice matched to GaAs. Since injection ratios depend exponentially on the band-gap difference [4] relatively small concentrations of Al (15-20%) are normally adequate to give freedom in emitter and base doping. Very large current gains can be achieved or traded for decreased emitter doping and increased base doping leading to lower capacitance and lower base resistance. The use of the heterojunction provides an extra degree of freedom to the designer in choosing doping profiles. Band-gap grading may also be used to provide field gradients to reduce transit times in the base.

There are, however, some disadvantages in having an AlGaAs emitter. These include increased off-set voltage and reduce carrier lifetime at the heterojunction itself [5].

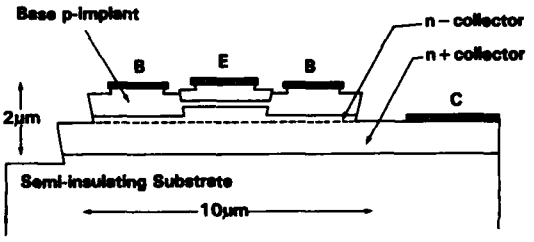


Figure 1 Plessey Process II Cross-section

2 THE PHYSICS OF BIPOLAR HETEROJUNCTION TRANSISTORS

Figure 1 shows a cross-section of the Plessey process II BHJT structure. It is an implanted base structure and will be taken as a typical geometry for our discussion of device physics. For convenience we can divide the problem into a one-dimensional structure following minority carriers vertically through the layers from emitter to collector and a two-dimensional network incorporating the lateral base geometry, extrinsic implant regions and the emitter, base and collector contact regions. Such a separation has its limitations since the device is really three-dimensional and the effects in the third dimension, illustrated by the SEM plan view in figure 2 due to via pad proximity etc. are not negligible. The approximations do however, prove to be effective in predicting measured results.

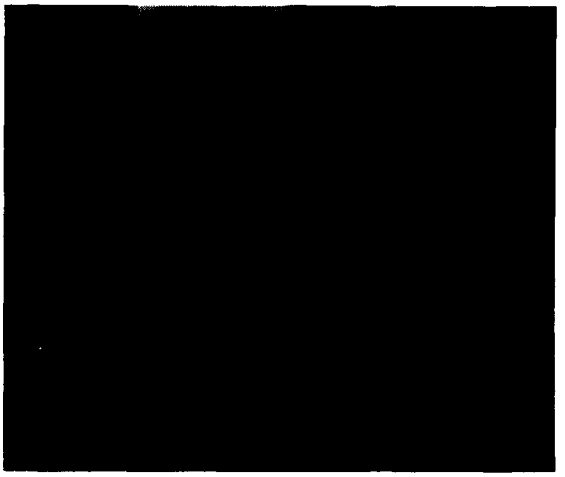


Figure 2 Plessey Process II Plan

2.1 One-dimensional Structure

Figure 3) shows a simulated band diagram for typical BHJT layer structure in one-dimension under zero applied bias [6]. The aluminium concentration has been deliberately graded with a parabolic position dependence in the region of the junction. This has the effect of removing the conduction band 'spike' brought about by the conduction band off-set. Much work has been performed on the effects of the conduction band spike on injection efficiency [4] hot electron effects [7] and enhanced recombination. It has been concluded that thermalisation is likely to be very rapid leading to no useful speed enhancement from hot electron effects and most designers choose to suppress the spike by grading to enhance injection efficiency.

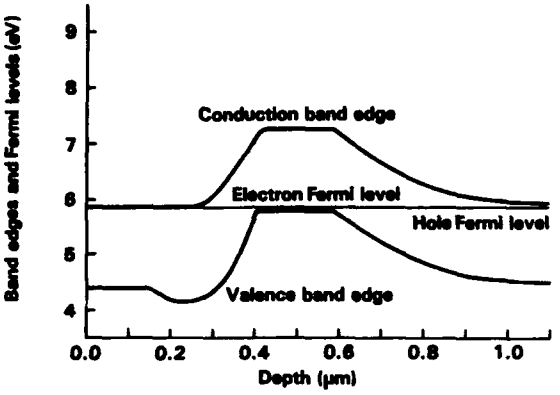


Figure 3 Simulated Band Diagram for Zero Bias

With a graded junction it is reasonable to approximate the injection physics by simplifying to the usual homojunction case with the built-in voltage modified by the net band off-set voltage. More detailed analyses are available [4] but the above simplification allows an important generalisation of the device models and permits the use of standard homojunction circuit simulation techniques such as SPICE [8] allowing rapid device and circuit design for BHJTs based on existing silicon BJT circuit families.

2.2 Two-dimensional Equivalent Circuit

Parasitics in the two-dimensional equivalent circuit slow down the practical BHJT device. Relatively large extrinsic capacitances (extrinsic base-collector) and even intrinsic (base-emitter and base-collector) capacitances are typical of devices currently reported. In addition, current state of the art contact resistances on GaAs are significantly higher than corresponding silicon technology. Geometries can be reduced to solve the parasitic capacitance problem but contact resistances become dominant. Emitter contact resistance can be very large and dominates the total emitter resistance although the use of a GaInAs layer will improve this as discussed later. Similarly, the large base contact resistance means that the total base resistance is dominated by extrinsic effects and the importance of the intrinsic distributed base resistance (dominant in silicon) is much reduced.

We conclude that the BHJT is describable by a modified BJT model with more emphasis placed on the extrinsic 'parasitic' network which can, quite effectively, be bolted on to the intrinsic 'one-dimensional' description.

3 MODELLING THE BHJT

It is vital that any programme of modelling should be both predictive and capable of relating basic design characteristics such as geometry and doping directly to circuit performance. Although optimisation and massaging of parameters to fit DC and RF measurements has its place when tuning up models on 'production' devices for use in advanced C.A.D., it is of very limited use at the research and design stage.

3.1 Parameter Extraction and Equivalent Circuit Development

We describe in this section a method we have developed to provide a complete modelling strategy for predictive design [9]. Parameters are extracted from basic geometry, layer specifications and test structure measurements. A variety of small and large signal equivalent circuits are generated and these are then used in simulation packages to calculate discrete transistor DC and RF characteristics and full logic circuit performance. Work is also under way to develop packages for large signal microwave simulation for power BHJTs based on methods developed for MESFETs [10]. The overall modelling scheme is displayed as a flow chart

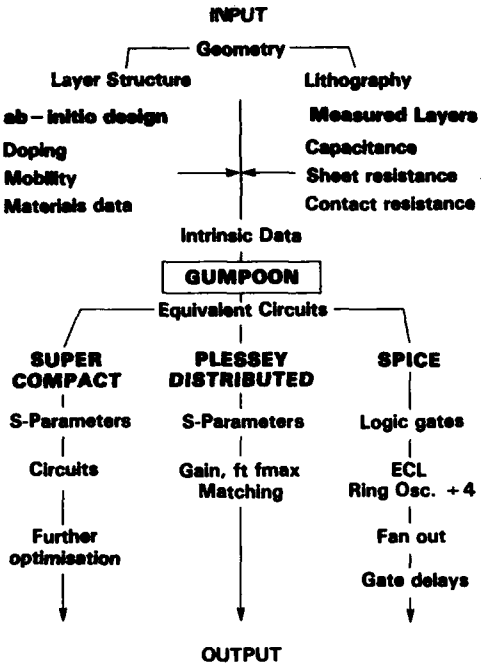


Figure 4 Modelling Flow Chart

in figure 4. Two inter-related routes are illustrated below the basic geometry inputs. For purely predictive modelling, emphasis is placed on the left hand 'ab-initio route'. Here the designer only specifies his required doping profiles and materials data from which much of the junction capacitance sheet resistivity and carrier mobility must be extracted. Once a layer has been grown and large geometry test structures processed more direct information on sheet resistivity junction capacitance and contact resistance can be supplied via the right hand 'measured layers' route. We stress, however, that these measurements are still at the fundamental level of layer and process parameters, actual transistor characterstics are not used at this stage. Indeed, measurement of layer capacitance is often a more reliable test of actual grown layer and doping structure than any direct attempt to measure doping profiles (by SIMS for example).

The small and large signal equivalent circuits are based on the Gummel and Poon and Ebers-Moll models [8]. The resultant circuits are appropriate for further analysis in various packages. Small signal analysis using 'SUPERCOMPACT' generates s-parameters, provides links into MMIC design and gives options for further optimisation once measured device characteristics are available. A small signal simulator has been written [9], which treats the lateral spread of the emitter, base and collector layers as three coupled active transmission lines terminated in the extrinsic impedances provided by the base and collector contacts. The basic circuit structure is shown schematically in figure 5. This distributed model is particularly suited to large power devices, and is an important improvement over lumped element transmission line models used for silicon BJTs.

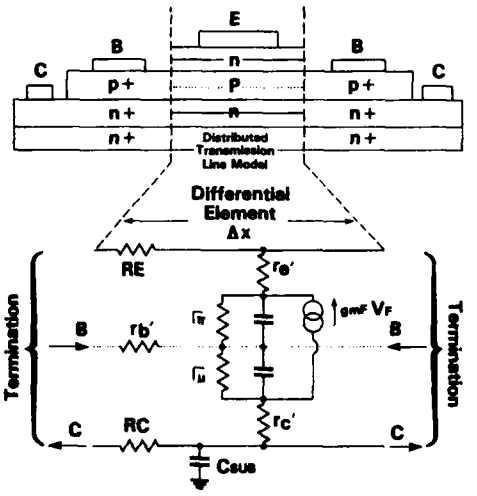


Figure 5 The Plessey Distributed Circuit Model Showing the Transmission Line Active Element

Finally, a large signal model is provided for SPICE simulations. Here the discrete transistor can be simulated at DC and RF and most importantly full logic circuits can be evaluated. This last activity is important since the performance of individual devices is critically dependent on their input and output matching. The only reliable way to determine the performance of a particular device design is to measure and simulate its performance in a full circuit. To this end we choose ECL divide by four circuits because they contain typical gate configurations used in ECL design.

3.2 Intrinsic Device Model

Inside the parameter extraction program the input data is processed using conventional device models. Junction and diffusion capacitances are calculated using the usual transistor theory [8] with the modifications for band offsets discussed in section 2 [4]. Parasitic resistances are calculated from the layer and contact dimensions together with the layer resistance and contact resistance either calculated from mobility data or entered directly as a measured parameter.

Transport of minority carriers is calculated on a drift and diffusion model in the base, allowing for any built-in fields provided by graded aluminium concentrations. In the collector, a saturated drift model is used. A more sophisticated description of the transport can be obtained from a finite difference numerical model [6] which takes a proper account of velocity field characteristics and the effect of the heterojunction. Either way the resultant transport is described in the equivalent circuit by an effective diffusion capacitance through the Gummel-Poon (GP) description [8].

4. Characterising the BHJT

Characterisation measurements for BJHT devices owe much to experiences on Si BJTs. However, by its very nature the BHJT is a high frequency device and requires full blown microwave techniques to measure at RF.

4.1 DC Characterisation

Standard DC measurements of Gummel Plots [8], DC and AC current gain, verses collector current, ideality and I-V characteristics are made for forward and reverse operation. These can be used to improve the basic DC model by specifying parasitic leakage effects on the gain and non ideal diode characteristics, both of which are difficult to quantify in a predictive way. We would highlight this area as being one requiring further work in order to close the loop on a completely predictive model.

4.2 RF Characterisation

As discussed in section 3 the most informative RF measurements are those made on full logic circuits. Such tests have been made by on wafer RF probing and then subsequently the selected circuits have been bonded into a purpose built RF jig for analysis with minimum parasitic effects.

5 RESULTS AND COMPARISON OF SIMULATIONS AND MEASUREMENTS

We report results obtained on the Plessey Process II devices and circuits with two different emitter sizes. The total emitter dimensions of the large device are 4 microns x 10 microns (metal contact width is 2 microns) and the smaller device 2.5 microns x 5 microns (metal contact width 1.5 microns). Calculations have been performed using

parameters appropriate to current processed devices which can be compared with measurements.

5.1 Discrete Transistors

Discrete device results are for an earlier batch of transistors to those quoted in the full circuit results below. In all cases a nominal 0.2 microns wide base layer doped at $4 \times 10^{18} cm^{-3}$ is assumed. Via pads, extrinsic to the device are isolated using an O_2 implant which leaves a residual parasitic capacitance. This is included in the model and can be reduced to a small value when predicting future device performance when a deep proton implant is implemented.

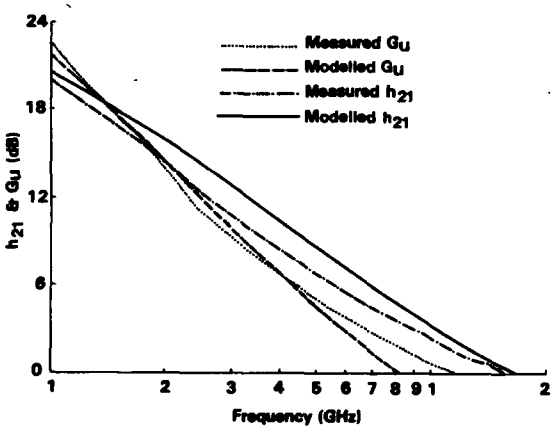


Figure 6 Measured and Modelled Gain for a 4μm Device

Figure 6 shows a comparison of the measured and modelled gain characteristics for a 4 micron device. It is stressed that *no subsequent 'optimisation' of the model has* been used to improve the fit. All the model parameters are as predicted from the known layer specification and device geometry. The agreement is very encouraging.

5.2 Logic Circuits

Both ECL ring oscillators and ECL divide by four circuits have been processed, measured and simulated. Ring oscillators functioned successfully under 5mA and 2mA tail currents. However, we found experimentally and confirmed theoretically, that higher order modes are easily excited. These confuse the interpretation of gate delays and cast doubt on some of the very fast results reported in the literature.

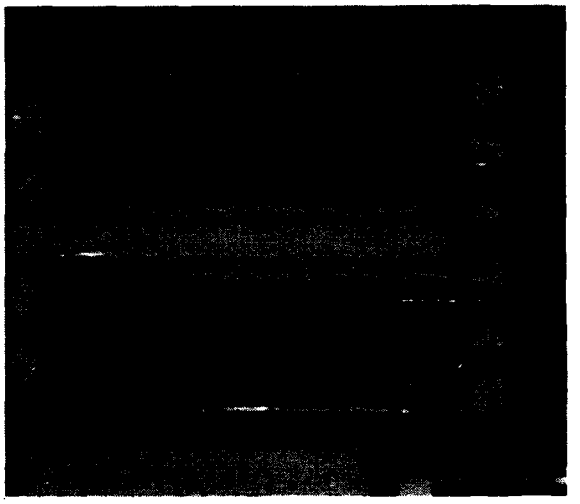


Figure 7 Micrograph of the Completed Divide by Four Die

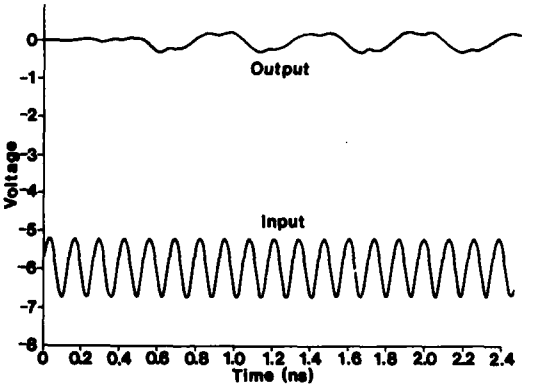


Figure 8 SPICE Prediction of Circuit Performance

Divide by four circuits have been fabricated in both 4 micron and 2.5 micron transistor variants. Figure 7 shows a photograph of a fabricated circuit. By taking a typical set of layer and process parameters a SPICE simulation predicted a maximum toggle frequency of 7.7GHz for 2.5 micron circuits. The simulated waveforms are shown in figure 8. Figure 9 shows the waveforms for a fabricated 2.5 micron divide by four circuit measured on wafer. The circuit toggled at up to 5.7GHz. The agreement is very satisfying particularly in the shape of the output waveform.



Figure 9 Measured Performance Showing a Divide by Four Circuit Toggling at 5.7GHz

Results for the 4 micron circuits on an earlier batch showed much closer agreement. In this case the predicted toggle frequency is 3GHz and measured is 3.3GHz. Of particular interest is the failure mode as the frequency is increased. The measured circuits evolve into a divide by six mode and the simulations predict exactly this failure. This adds to our confidence in the predictive capability of the model. Power consumption can be very significant since in order to maintain the gain all the transistors run at between 2.5mA and 5mA. Measured power consumption for the 2.5μm divider is about 1 watt at present but this will be significantly reduced when smaller geometries are used. Chang et al [11] report a 2μm CML divide by four circuit toggling at 11GHz with a power consumption of 315mW.

5.3 Effects of Process Variation and Improvement

Having validated the model against measurement it is possible to identify which parameters are most critical and require improvement and which are less sensitive and can be traded.

In table 1 we catalogue the effect of taking our basic parameter set (used in the simulation of figure 8) and varying a limited selection of parameters, changing just one in each case. Emitter and collector contact resistances are seen to be significant and in the case of the collector a significant increase in circuit output swing was also predicted.

The effect of reduced base thickness is interesting since it increases the intrinsic base resistance and trades this for shorter transit time. The net effect is a significant increase in divider toggle frequency showing clearly that the intrinsic base resistance is not critical. If all the critical parameters are set at our expected achievable values (as seen in our development experiments) the predicted best operation for 2.5μm devices is a divide by four toggle frequency of 11GHz in the case of using O_2 as the via pad isolation and 23GHz using a total isolation procedure with protons.

6 BHJTs: ALTERNATIVE STRUCTURE

Various alternative structures have been developed throughout the world. The basic layer structures feature both single heterojunction [11] and double heterojunction [5] variants. Inclusion of a second heterojunction at the base-collector interface

TABLE 1

Parameter Varied	Emitter Contact Ωcm^2	Collector Contact Ωcm^2	Base Width μm	fmax GHz	f4* GHz
Basic Set	5×10^{-6}	2×10^{-5}	0.2	13	7.7
Emitter Contact	2×10^{-6}	2×10^{-5}	0.2	15	8.9
Collector Contact	5×10^{-6}	2×10^{-6}	0.2	15	8.6
Base Width	5×10^{-6}	2×10^{-5}	0.1	15	8.7

* maximum toggle frequency for a divide by four circuit

is intended to reduce offset voltage, offer flexibility in operating transistors in forward and reverse mode and reduce hole injection into the collector (particularly important for saturating logics such as I^2L but not critical for ECL).

Consideration has been given to the advantages of an inverted structure in which the collector layer is grown on top [12]. This has the important advantage of reducing collector capacitance although at the expense of increased collector resistance (small area contact) and increased emitter-base capacitance.

Contact to the base layer is normally achieved either by direct implantation [9] or by a mesa etch to the base [11]. Implantation has the advantage of retaining planarity of surface and also lends itself to easier self alignment structures. A possible disadvantage is the creation of damage leading to leakage paths and parasitic diodes which reduce current gain. The direct etch to the base layer has the disadvantage that the base layer is very thin (0.05-0.3 microns) and such an etch has to be very accurate, also ohmic contacts have to be formed onto this layer.

Reduction in collector capacitance has been achieved in the direct etch structure by using a deep proton implant through the base contact region to isolate some of the extrinsic base-collector area [13]. Self alignment techniques also reduce this extrinsic area [13] and address the problem of emitter contact resistance in small geometry structures by extending the emitter contact up to and beyond the edges of the emitter crystal [13].

7 BHJTs: WORLD POSITION

At the time of writing the world best position for BHJT circuit performance can be summarized by table 2.
These results are in line with our model prediction which suggests that the Plessey 2.5μm process should run with divide by 4 circuits toggling at well in excess of 11GHz.

8 INTEGRATED CIRCUIT PROCESSES

In digital logic both silicon and GaAs techniques are now competing for gate delays much less than 30ps. In GaAs systems, MESFET logic offers quite high speed with relatively low power consumption. MESFET BFL divider

TABLE 2

NTT [13]	Self aligned with proton implanted regions 2 x 5μm² ECL
ft = 45GHz fmax = 65GHz	
f4* = 13.7GHz	
Rockwell [11]	Self aligned base contact with proton implanted regions 2.5 x 4.5μm²
ft = 45GHz	
f4* = 11GHz (CML)	

* Maximum toggle frequency for a divide by 4 circuit

circuits have been reported, toggling at up to 17.9GHz [14], dissipating 657mW. MESFETs do however suffer from threshold uniformity and fan out problems (low current drive). Some improvement in uniformity has been reported by the use of HEMT devices but is still a problem. BJHTs have by comparison displayed excellent standard deviation in threshold voltage (2.5mV across a one inch square wafer reported by Chang et al [11]) and have the intrinsic advantage of substantial current gain for good fan-out. The MESFET result featured a 0.25μm gate length, which is likely to cause yield problems whereas the BHJT is predicted to be capable of speeds in excess of 18GHz with critical dimensions greater than 1μm.

Super self aligned Si BJT circuits have been fabricated [1] with 30ps gate delays and nearly 10GHz toggle frequency for divider circuits and 7K gate complexity has been reported by NTT. Similar results have been reported by Hitachi using a variant process. However, the NTT result features a 0.35 micron wide emitter which compares with the current 2-micron wide BHJT emitter toggling at the same frequency [11, 12]. Self aligned BHJT designs which reduce the emitter size substantially will give outstanding improvements in speed and power consumption and lead to increased integration. It is clear that the Si devices are stretched to the limit whereas the BHJT has plenty of room for improvement.

In microwave power the race is again between the BHJT, the Si BJT and the GaAs MESFET. At X band both Si [4] and MESFET [10] have shown output powers from 1-2W. No measured results are available for the BHJT but its intrinsically higher speed (fmax ~65GHz as quoted in section 7), should give it the edge over Si bipolar. The fact that it is a bipolar device means that it can carry a higher current density than the MESFET and so, at least in pulsed mode, much higher powers at higher frequencies are possible.

9 CONCLUSIONS AND A LOOK INTO THE FUTURE

We conclude that the BHJT has a solid place in the fields of high speed logic and microwave power. It has already obtained a slight edge over its rivals in divider performance and is developing smaller geometries and self aligned techniques which will ensure its dominance. Microwave power performance is as yet untried but the extra power of the bipolar concept combined with the speed of GaAs must make it a threat to the MESFET at high frequency and the Si BJT at intermediate frequencies.

Many options exist for improved technology, like the use of refractory contacts and GaInAs ohmic contact layers. Higher speed, lower power and higher integration will come from self aligned structures with smaller geometries (one micron or less) and improved contact resistance. New materials such as GaInAs in the base of a GaAs device offer improved performance [15]. The addition of indium gives the designer freedom to vary the band gap below

that of GaAs. Graded bases are therefore possible leading to built-in fields without detracting from the injection efficiency which is maintained by the AlGaAs emitter. It may even be feasible to make a GaAs emitter with a GaInAs base. The reduced overall band-gap will lower the operating voltages and improve circuit performance and design tolerance. The strain introduced by the indium is thought to be acceptable (devices have been successfully fabricated in this system [15]) and may even lead to a strained layer effect on the valence band reducing the hole mass and improving mobility! The possibility of using InP based structures with GaInAsP bases [16] offers improved performance and possible optoelectronic integration.

There now exists some interesting speculation on possible new structures such as induced base transistors [17] and quantum well bases [18]. The induced base device offers an improvement on the metal base transistor since it uses a 2D electron gas as a base which provides low base resistance but allows hot electrons to penetrate to the collector with high gain. A quantum well in the base is a useful physics tool since it allows for the controlled injection of carriers into the well without distorting its shape. A possible three level logic device has also been proposed [18]. These are but two examples of the exciting structures that can be envisaged when the concepts of band-gap engineering are used. Future transistor designs must exploit this to the full.

ACKNOWLEDGEMENTS

The author acknowledges the work of R. C. Goodfellow, R. C. Hayes, P. J. Topham, C. G. Eddison, C. J. Greenwood, I. Goodridge, D. V. A. Benn and J. G. Metcalfe. Part of this work was supported by the Procurement Executive, Ministry of Defence, U.K. (DCVD), sponsored from R.S.R.E. Malvern and part was supported under ESPRIT, project No. 971.

REFERENCES

[1]. S. Konaka, Y. Yamamoto and T. Sakai, IEEE ED 33, 526, 1986

[2] G. W. Schreyer, Microwave Journal, 24, 63, 1981

[3] H. T. Yuan and Y. S. Wu, Proc. IEEE Int. Microwave Sympt., Ottowa, 378, 1978

[4] A. Marty, G. Rey and J. P. Bailbe, S.S. Elec., 22, 549, 1979

[5] S. Tiwari, Proc. IEDM 1986, 262

[6] E. Azoff, PhD. Thesis, University of Sheffield, UK., 1987

[7] T. Ishibashi, H. Ito and T. Sugata, Proc. Int. Symp., GaAs and Related Compounds, Biarritz, 1984, 593

[8] I. Getreu, 'Modelling the Bipolar Transistor', Tektronix Inc. Beaverton, Oregon, 1979

[9] A. J. Holden, C. G. Eddison. J. G. Metcalfe and R. C. Hayes, Elec. Lett., 22, 815, 1986

[10] A. J. Holden, B. T. Debney, J. P. King, J. G. Metcalfe and C. H. Oxley, IEE Proc. 133, Pt. H., 399, 1986

[11] M. F. Chang, P. M. Asbeck, K. C. Wang, G. J. Sullivan and D. L. Miller, Elec. Lett., 22, 1173, 1986

[12] K. Morizuka, T. Nozu, K. Tsuda, M. Azuma, Elec. Lett., 22, 315, 1986

[13] O. Nakajima, K. Nagata, Y. Yamauchi, H. Ito and T. Ishibashi, Proc. IEDM 1986, 266

[14] J. F. Jensen, L. G. Salmon, D. S. Deakins and M. J. Delaney, Proc. IEDM 1986, 476

[15] P. M. Enquist, L. R. Ramberg, F. E. Najjar, W. J. Schaff and L. F. Eastman, App. Phys. Lett., 49, 179, 1986

[16] H. Krautle, Elec. Lett., 22, 1192, 1986

[17] C. Y. Chang, W. C. Liu, M. S. Jame, Y. H. Wang, S. Luryi and S. M. Sze, IEEE EDL 7, 497, 1986

[18] F. Capasso, S. Sen, A. C. Gossard, A. L. Hutchinson and. J. H. English, IEEE EDL 7, 573, 1986

Session A2.3

Ultrafast Bipolar III

Chairman: L. Treitinger

Tuesday, September 15, 1987

CHARACTERIZATION OF OPTIMIZATION OF BIPOLAR TECHNOLOGIES BY MEANS OF HIGH SPEED CIRCUIT DESIGN

Dr. Wilhelm Wilhelm

Siemens AG, WIS P SL 2
Balanstraße 73, D-8000 München 80

The speed of bipolar technologies is limited by two time constants. One of these time constants is caused by parasitic effects, the other one by the base transit time. In this paper an attempt is made to weigh both figures for optimum performance at different circuit designs.

1. INTRODUCTION

The limits of performance data of bipolar integrated circuits are determined by the dynamic parameters of the technology. A commonly used figure of merit is the expression: [1]
$F = [FT/(2\pi \times RB \times CP)]^{1/2}$, where FT is the transit time, RB the base resistance and CP stands for parasitic capacitances of the bipolar transistor. However, this figure can not be applied to all types of circuits because the optimum ratio of
$\tau B = 1/(2\pi FT)$ to $\tau P = RB \times CP$ in addition to the value of F must also be considered.

By examining three specific cases of bipolar design, this paper will show the relationship of the time constant quotient to the optimized performance in these design areas.

2. TECHNOLOGY TIME CONSTANTS

The two basic parameters, τB and τP, can be calculated for a bipolar technology with fixed design rules. The first parameter, τB, is dependent upon the charge needed in the base region to switch on or off the bipolar transistor. The time constant τB is is more dependent on the vertical doping profile than on the lateral dimensions of the transistor. The second time constant, τP, mainly consists of parasitic capacitances which are charged by parasitic resistors formed by the base resistor in the bipolar transistor. These parameters are proportional to the lateral dimensions of the transistor. As the transistor is reduced in size, the time constant is also reduced due to the smaller area. However, for fixed design rules, the time constant τP is almost independent of the emitter length: As the capacitance increases with emitter length, the base resistance also decreases. Fig. 1 shows the time constant for collector-base, base-emitter and the collector-substrate as a function emitter length.

Normalized to the collector-base time factor one gets roughly:

$\tau C = \tau P$
$\tau E = 2 \times \tau P$
$\tau S = 5 \times \tau P$

Former and older technologies show similar or nearly equal relationships but tend to use a factor of 3 or 4 for the substrate value.

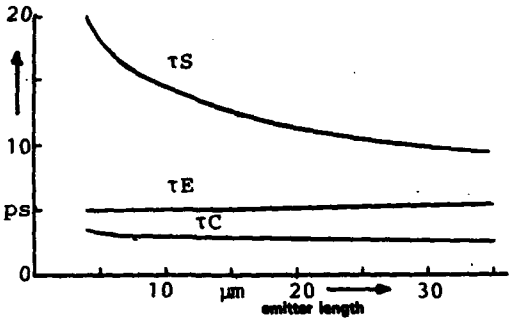


FIGURE 1

Parasitic time constant versus emitter length

τS = RB x CJS
τE = RB x CJC
τC = RR x CJC

3. CIRCUIT DESIGNS

Three typical circuits are analyzed on terms of parasitic and transit time limits.

3.1. Bipolar optical receiver

The first one is a typical analog example: A high speed transimpedance amplifier. The figure which in this case has to be optimized is the equivalent input noise current. This parameter limits the sensitivity of the fiber optic receiver. The noise is mainly generated by the input transistor of the receiver. Noise sources at high frequencies are generated by the base-resistor and the shot noise of the collector. The equivalent input noise can be minimized by optimizing the operating current and the transistor size. [2]

At low supply current the shot noise dominates because of small Transconductance. At high currents the increasing diffusion capacitance induces higher equivalent input noise. Similar effects are observed by varying the size of the transistor. Smaller transistors generate higher thermal base resistor noise, but in very large transistors the parasitic capacitance dominates and the shot and base noise translated to the input is increased. At the optimum operating current and transistor size, we have an expression which contains only the two parameters τB and τP.

$$\overline{I}_{R^2} = 20\ KT\ FB^3.CD/[(\tau P.\tau B)^{1/2}.F1]$$

where FB = Bit frequency
CD = Capacitance of the optical device
K = Balzmann constant
T = absolute temperature
$F1 = R+1/R + (3/R + R^3)^{1/2}$
where $R = (\tau_p/\tau_B)^{1/2}$

The function F1 plotted in Fig. 2 shows a relatively small mimum at τP/τB = 0.7.

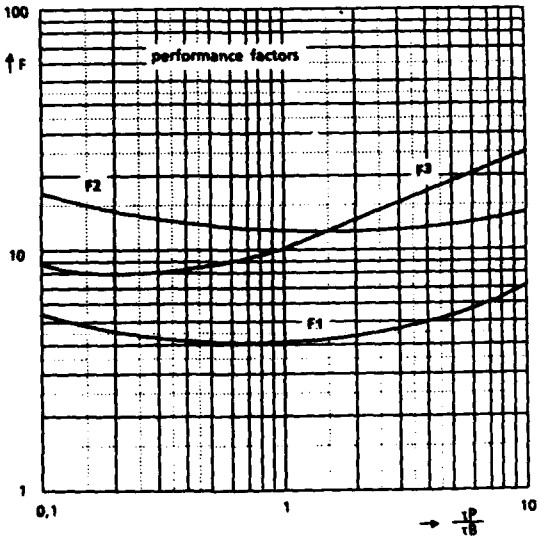


FIGURE 2

Performance factors for 3 different circuit designs

3.2. CML Gates

For the second example, a CML logic design is chosen. This design is used for LSI or VLSI circuits. The figure of merit is the highest available speed. The optimum will be found at a supply current which is able to handle parasitics but is small enough not paying penalty by high injection effects and base resistance limitations. At this minimum the propagation delay (fan in = fan out = 2) depends on the two time constants only.

$TPD = (\tau P \times \tau B)^{1/2} \times F2\ (\tau P/\tau B)$

F2 is plotted in Fig. 2.

It shows a relatively flat minimum at $\tau P/\tau B = 1$. However, at smaller $\tau P/\tau B$ the power consumption gets smaller. Therefore the optimal value should be in the range of .2 to .5 rather than 1.

3.3. Telecom circuits for high data rates

The third example is a digital circuit for serial data processing which enables processing of highest data rates. Not the delay time but the rise and fall time of the serial signal has to be optimized. Differential logic is used in this case with emitter follower driver. Due to the differential nature of the signal the logic swing is small and ranges between .2 and .3V. The optimum load resistance is smaller than in the previous cases. The resulting data rate (DR) in terms of B and P looks like:

$DR = 1/[(\tau P \times \tau B)]^{1/2} F2\ (\tau P/\tau B)$

The function F2 is plotted in Fig.2. The minimum tends to be more at the low τP side of the plot.

4. CONCLUSION

In this paper an attempt is made to relate the inherent time constant on the dynamic parameters to each circuit design. Due to the progress of the MOS technology, which is able to produce channel length of 1 µm and even below this value, the tendency is obvious to reduce parasitics by scaled down transistors, by self aligning techniques and improved isolation techniques like trench isolation. However, no or little effort is made to increase simultaneously the cut off frequency of the transistor to get an optimum relationship between both time constants. In modern technologies this ratio ranges between 0.05 to 0.02 which is far away from optimum for most circuit designs. There may be other than performance considerations which result to such adequate ratios. These could be for instance yield or available fabrication equipment or other limitations.

REFERENCES

[1] S.K. Ghandi, The Theory and Practice of Microelectronics, John Wiley and Sons, Inc., New York, 1968

[2] K. Panzer, K. Schrödinger, W. Wilhelm, Digital Optical Transmission System for Industrial Use with Data Rates of up to 200 Mbit/s, Siemens Components 5/86

Physical modelling problems of ultrafast silicon bipolar transistors

by

H.C. de Graaff and G.A.M. Hurkx

Philips Research Laboratories
5600 JA Eindhoven, The Netherlands

1. Introduction

This paper deals with recent developments in physical models that are of special importance for the ultrafast, very high frequency behaviour of vertical bipolar npn transistors: minority carrier mobility, bandgap narrowing, hot carrier transport in collector epilayers and junction sidewall effects.

Transistors with a high cut-off frequency f_T (10-20 GHz) and a low r_bC_{cb} product ($\leq$4 ps) are discussed, i.e. devices with short transit times ($\sim$10 ps), low base and emitter resistances and low capacitances. That in turn requires shallow junctions ($x_j \approx 0.10 - 0.15$ μm) and reduced lateral dimensions, making sidewall effects increasingly important.

2. Transit times

The maximum value of f_T is determined by the total transit time of the emitter, base and collector regions: $1/2\,\pi\,f_T = \tau_e + \tau_b + \tau_c$. Simple approximations can be deduced for each contribution [1]:

$$\tau_e \approx \frac{1}{h_{FE}} \frac{W_e}{D_p}\left(\frac{1}{2}W_e + \frac{D_p}{s} + \frac{N_0}{dN/dx}\right), \quad (1)$$

$$\tau_b \approx \frac{W_b^2}{\eta D_n} + \frac{W_b}{v_s} \quad \text{and} \quad (2)$$

$$\tau_c \approx \frac{W_{epi}}{2\,v_s} + r_{epi}C_{cb}, \quad (3)$$

where W_e and W_b are the thicknesses of the neutral emitter and base regions, s is the recombination velocity at the contact, N_0 the dope concentration at the contact and dN/dx the steepness of the profile at the junction, v_s is the saturated drift velocity and η is the built-in field factor (η = 5-8 for these transistors). The other symbols have their usual meaning.

Typical values are $D_n \approx 10$ cm^2/s, $D_p \approx 2$ cm^2/s and $v_s = 10^7$ cm/s. With $W_b \approx 0.10$-0.15 μm, τ_b becomes about 3 ps; if we take $W_{epi} = 0.4$ μm, then $\tau_c \approx 2.5$ ps, but this may be doubled when the collector charging time $r_{epi}\,C_{cb}$ increases due to hot carrier flow in the epilayer. For an emitter with a polysilicon contact we have e.g. $h_{FE} = 100$, $s = (0.5\text{-}1.0) * 10^5$ cm/s and $No/(dN/dx) = 10^{-5}$ cm, so $\tau_e \approx 2$ to 3 ps. An aluminium contact has $s \approx 3 * 10^5$ cm/s, but it lowers h_{FE} and usually dN/dx is also decreased, so τ_e may grow to 10 ps.

The examples given show that:

- all regions contribute significantly to the total transit time;
- the value of the current gain (h_{FE}) must remain relatively high (>50);
- the doping profiles near the e-b (dN/dx) and the c-b junctions ($r_{epi}C_{cb}$) are crucial.

3. Bandgap narrowing, mobility and recombination

The h_{FE} depends on the bandgap narrowing and minority carrier mobilities in the emitter and base and on the minority carrier lifetime (τ_p) and the contact recombination velocity(s) in the emitter.

The minority carrier mobility as a function of the dope concentration (N) is written as

$$\mu = \mu_{min} + \frac{\mu_{max} - \mu_{min}}{1 + (N/N_{ref})^a}. \quad (4)$$

The mobility parameters in eq. (4) are given in Table 1.

For bandgap narrowing we can mention the following models:

- Slotboom and de Graaff (SdG) [2] give the bandgap narrowing as $\Delta V_{gap} = 9 * 10^{-3} (F + \sqrt{F^2 + 0.5})$ volts, with $F = \ln (N/1 * 10^{17})$ for p-type as well as n-type material.
- del Alamo, Swirhun and Swanson (dASS) [3] found in n-type material $\Delta V_{gap} = 18.7 * 10^{-3} \ln (N/7 * 10^{17})$ volts.
- In p-type material Swirhun, Kwark and Swanson (SKS) [4] found the same ΔV_{gap} as in the SdG-model.
- Bennett (B) [5] took for the minority carrier mobilities at high doping levels ($N \sim 10^{20}$ cm^{-3}) $\mu_p \approx 3.3 * \mu_{p\,maj.}$ and $\mu_n \approx 1.4 * \mu_{n\,maj.}$. At these doping levels the ΔV_{gap} is given by

$$\Delta V_{gap} = 2 \frac{kT}{q} \ln\left\{1 + 1.90 \exp\left(- \frac{N_D - 2.4 * 10^{20}}{1.68 * 10^{20}}\right)^3\right\}.$$

Table 1

	μ_{max} (cm^2/Vs)	μ_{min} (cm^2/Vs)	N_{ref} (cm^{-3})	a	remark
SdG	468	49,7	$1.6 * 10^{17}$	0.7	n-type
	1360	92	$1.3 * 10^{18}$	0.91	p-type
dASS	500	130	$8 * 10^{17}$	1.25	n-type
SKS	1412	232	$8 * 10^{16}$	0.9	p-type

The base Gummel number (G_b) determines the collector saturation current density and is defined as

$$G_b = \int_0^{W_b} \frac{N_A(x)}{D_n(x)} \exp(-q\Delta V_{gap}(x)/kT)\,dx$$

$$= \left\langle \frac{\exp(-q\Delta V_{gap}/kT)}{qD_n} \right\rangle \int_0^{W_b} qN_A\,dx$$

With $\int_0^{W_b} N_A dx = Q_{bo}$ we get

$$\langle \mu_n \exp(q\Delta V_{gap}/kT) \rangle = \frac{Q_{bo}}{kTG_b} \qquad (5)$$

Q_{bo} is determined from measurements of the Early effect, G_b follows from measured collector saturation currents. By taking different geometries we have separated bulk and sidewall components. Fig. 1 shows the results for four different processes, compared with the SdG and SKS models and the case $\Delta V_{gap} \equiv 0$. The conclusion is that the existing bandgap models for the base region overestimate the collector current by some 30 to 50 percent. In most cases the product of mobility and bandgap narrowing in the base can be approximated by a constant (= 800 ± 100 cm^2/vs).

For the emitter Gummel number (Ge) not only ΔV_{gap} and D_p, but also the lifetime (τ_p) and the recombination velocity (s) are important [6]:

$$G_e \approx \left(0.85\,W_e + \frac{D_p(0)}{s}\right) g(0) \left\{\frac{N_{ii}}{D_p(0)} \exp(-q\Delta V_{gap}/kT)\right\} \qquad (6a)$$

$$g(0) = \frac{J_p(0)}{J_p(W_e)} \approx \frac{\tau_p(x_m)}{\tau_p(x_m) + W_e/4\,D_p(0) + \frac{1}{3}W_e/s}. \qquad (6b)$$

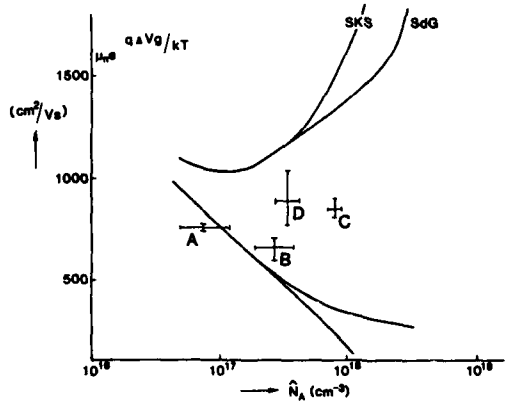


Fig. 1. Product of mobility and bandgap narrowing versus max. dope concentration for p-type base regions. Fully drawn lines are model predictions, the lower ones for the case ΔVgap = 0 and different mobilities.

$\tau_p(x_m)$ is the minority carrier lifetime at the spot of maximum recombination; the contact is at x = 0.
From measured values of G_e we have calculated $\tau_p(x_m)$ via g(0), using the SdG, dAS and B models for minor-

ity carrier mobility and bandgap narrowing. The three resulting sets of τ_p values are compared with experimental values from the literature in fig. 2. The dASS model gives the best results, the B model leads to τ_p values that are one order of magnitude too low.

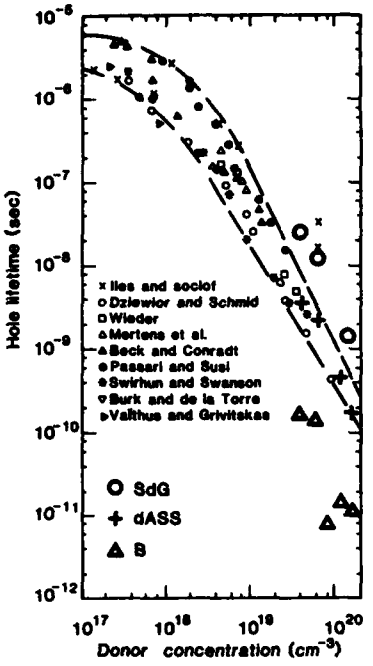


Fig. 2. Minority carrier lifetimes in n^+ emitters, from literature and from measurements using three different models.

4. Hot carriers in collector epilayers

For homogeneous samples in the static situation we have $\vec{J} = q\mu n \vec{F}$ and $\vec{F} \cdot \vec{J} = \frac{W - W_0}{\tau_W}$, where W is the carrier energy density and τ_w the energy relaxation time. Introducing an (effective) carrier temperature (T_c) by means of $W = 3/2k\ nT_c$ results in

$$T_c - T_0 = \frac{2}{3}\frac{q}{k}\tau_w \mu F^2 \qquad (7)$$

- Cook [7] assumes that τ_w = constant = 0.4 ps and $\mu = \mu_0 (1 + \mu_0 F/v_s)^{-1}$. Substitution in eq. (7) gives $T_c(F)$ as shown in fig. 3.
- Baccarani et al. [8] start with a constant diffusivity and put $T_c/T_0 = \sqrt{1 + \left(\frac{\mu_0 F}{v_s}\right)^2}$, resulting in an energy-dependent τ_w.
- Hänsch et al. [9] take again τ_w as constant and put $\mu = \mu_0 / \left\{1 + \frac{3}{2}\frac{k}{q}\frac{\mu_0}{\tau_w v_s^2}(T_c - T_0)\right\}$.

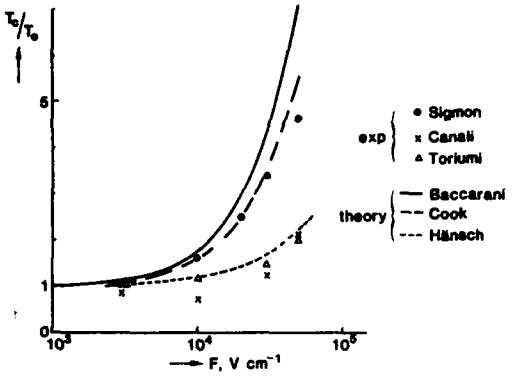


Fig. 3. Carrier temperature (T_c) versus electric field (F): experimental and theoretical. Cook uses τ_w = 0.4 ps, for the Hänsch model τ_w = 0.1 ps is taken.

These three models for T_c (F) are compared in fig. 3 with experimental values obtained from D(F) and μ(F) measurements [10, 11] : $T_c = \frac{qD(F)}{k\mu(F)}$, or from a spectral analysis of luminescence effects in MOS-transistors [12].

In the collector epilayer field strengths around 10^5 V/cm may occur, but neither theory nor experiments can yet tell us what carrier temperatures we can expect. Velocity overshoot will probably occur, but only over short distances of less than 0.1 μm [8]. In our opinion hot carrier flow in collector epilayers is still adequately described by the critical current density $J_{hc} = q\ N_{epi}\ v_s$; it gives an extra voltage drop in the collector epilayer and increases the collector charging time $r_{epi}\ C_{cb}$.

5. Emitter sidewall effects

Fig. 4 shows the results of 2-D simulations of the base and collector current densities crossing the e-b junction [13]. From such simulations we learn that the base sidewall current strongly depends on the ratio of the distance between contact and junction edge and the minority carrier diffusion length (Δ/L_p), and on the recombination velocity (s) at the contact.

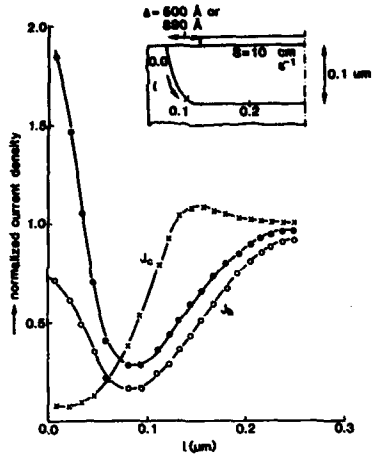


Fig. 4. Calculated base and collector current densities crossing the e-b junction. The highest J_b belongs to Δ = 500 Å.

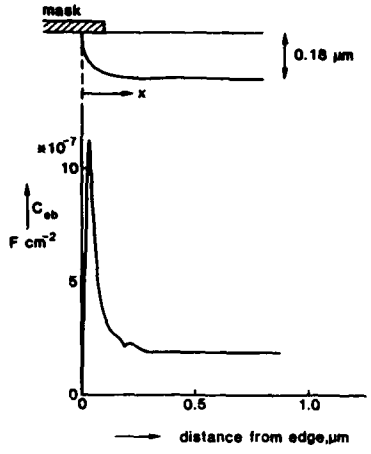


Fig. 6. Calculated depletion capacitance density along the e-b junction at V_{bc} = 0 volt. x_{jc} = 0.18 μm.

For $\Delta/Lp \gg 1$ and low s-values (use of spacers and polysilicon contacts) we even get that h_{FE} increases for smaller emitters (see fig. 5), at the expense, of course, of larger series resistances.

Fig. 6 shows the sidewall and bulk contributions of the e-b depletion capacitance.

The sidewall effects are clearly perceptible over distances of about 0.2 μm for both currents and the capacitance. So a separation between bulk and sidewall components only make sense for emitters wider than 0.4 μm.

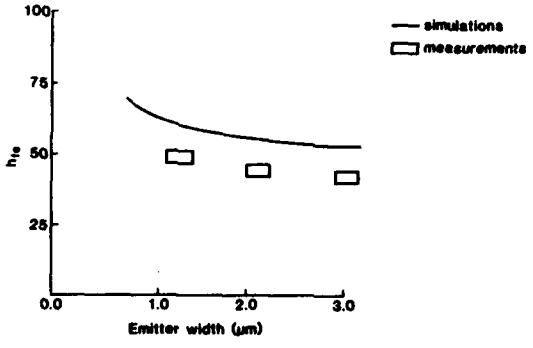


Fig. 5. Current gain (h_{FE}) versus emitter width. x_{jc} = 0.18 μm, x_{jc} = 0.40 μm. Al contact with Δ = 0.2 μm spacer width.

6. Conclusions

The existing models of bandgap narrowing, mobility and lifetime, when confronted with measurements, need some modifications. The hot carrier situation is rather confused, but for bipolar transistors the concept of saturated drift velocity is still useful. Sidewall effects become increasingly important with submicron emitters.

References

1) J.H. van den Biessen, Solid-St. Electr. **29,** 529 (1986).
2) J.W. Slotboom, H.C. de Graaff, Solid-St. Electr. **19,** 857 (1976).
3) J. del Alamo, S. Swirhun, R.M. Swanson, Solid St. Electr. **28,** 47 (1985).
4) S. Swirhun, Y.H. Kwark, R.M. Swanson, IEDM Techn. Digest **24** (1986).
5) H.S. Bennet, IEEE Trans. **ED-30,** 920 (1983).
6) H.C. de Graaff, J.W. Slotboom, A. Schmitz, Solid-St. Electr. **20,** 515 (1977).
7) R.K. Cook, IEEE Trans. **ED-30,** 1103 (1983).
8) G. Baccarani, M.R. Wordeman, Solid-St. Electr. **28,** 407 (1985).
9) W. Hänsch, M. Miura-Mattausch, J. Appl. Phys. **60,** 650 (1986).
10) C. Canali et al., Appl. Phys. Letrs. **27,** 278 (1975).
11) T.W. Sigmon, J.F. Gibbons, Appl. Phys. Ltr. **15,** 320 (1969).
12) A. Toriumi et al., SSD 86-13, 33 (1986).
13) G.A.M. Hurkx, IEEE Trans. **ED-34,** accepted for publication.

AN ADVANCED BIPOLAR PROCESS USING TRENCH ISOLATION AND POLYSILICON EMITTER FOR HIGH SPEED VLSI

M. ROCHE, G. BOREL, JL. IMBERT, D. THOMAS, L. FRITSCH, D. CELI
THOMSON SEMICONDUCTEURS BP 200 38522 Saint-EGREVE Cedex FRANCE

P. HUNT, PLESSEY Research LtD, Allen Clark Research Centre
Caswell Towcester, Northants NN129Q U.K.

A. HEFNER, TELEFUNKEN Electronic Gmbh, Therensienstrasse
D1700, Heilbronn, Postfach 1109, RFA

1. INTRODUCTION

Super computers and communication systems require very high speed circuits with a large degree of complexity such as signal processors, gate arrays, RAM'S, PROM'S etc... Bipolar technology, always two to four times faster than MOS technology is on the move again, now that it has overcome the scaling problems encountered with the MOS structures. As far as physical mecanisms are concerned no real limitations are foreseen at present. To achieve the upmost performances with bipolar technologies in the micron and submicron ranges one has to deal primarily with parasitic components. The only active part of the transistor is the one underneath the emitter. All the other regions are parasitic and decrease the performances. Their impact has to be reduced. To go in this way the dielectric filled trench isolation combined with self-aligned emitter base structure is the most promising solution for the next five years [1] [2].

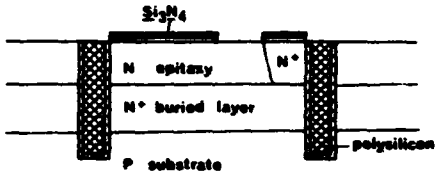
Today a rather limited number of real applications come into sight. The reasons are first of all the difficulties to fill perfectly the trenches without holes or crevices and secondly the defects generation at the corners of the trenches. This aspect is particularly important in the case of walled emitter and is one of the key issue of this paper. Concerning the self-aligned structure several schemes have been proposed up to now. Undoubtely the best solution consists of two levels of polysilicon closely separated by an oxide spacer, one level acting as the base electrode and the other level as the emitter electrode [3]. As this approach involves several RIE steps it does not guarantee against the damages induced during base polysilicon etching. In order to separate the variables for a clear understanding of the trench isolation problems we report the results concerning one polysilicon level.

2. DEVICE STRUCTURE AND FABRICATION STEPS

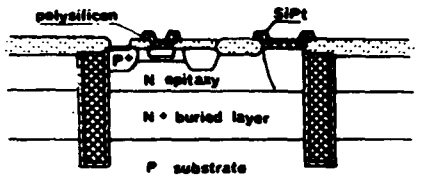
In the first phase the process has been set up using one micron design rules. The photolithography steps were realized with the CENSOR SRA 9535. We found that alignment and focus based on the procedure using phase contrast in bright field were very sensitive to the thicknesses of the different layers. To overcome these difficulties we developed special scattering alignment marks and new coating resist process with very low pile up. For submicron lines a reversal resist process using the AZ5214E product was studied.

The concepts of the process sequence are explained in Figure 1 (A)-(D) which illustrates cross-sectional views of the NPN transistor.

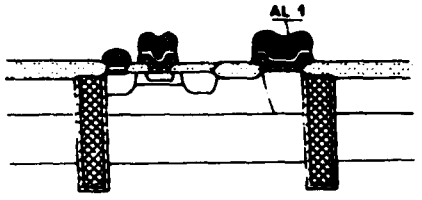
The process starts with a tri-step (high temperature-low temperature-high temperature) intrinsic gettering procedure performed on p type CZ pulled <100> orientated silicon. A full sheet arsenic burried layer is implanted and annealed followed by the 0.9µm thick epitaxy deposition. Special requirements for deep trench isolation have been described [4] [5].



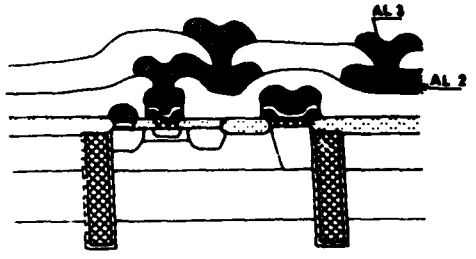
a) ACTIVE AREA DEFINITION



b) POLYCIDE FORMATION



c) FIRST METAL INTERCONNECT



d) THIRD METAL INTERCONNECT

. FIGURE 1 - Fabrication steps for NPN transistor

A conventional process was developed that combines :

- reactive ion etching employing CCl_4 gas, where stacked layers of thermal silicon dioxide, LPCVD nitride and PECVD oxide are used as the etching mask
- oxidation of the trench sidewall to provide dielectric isolation
- boron ions implantation in the bottom of the trench for channel stopper creation
- conformal undoped polysilicon deposition to refill the grooves
- planarisation of this polysilicon layer
- oxidation for both field oxide and polysilicon capping formation.

Two specific points have been studied :

- low defect generation

. chemical sidewalls cleaning prevents contaminated and/or damaged silicon layer during the RIE to interface

. rounding the trench bottom by controlling the redeposition rate of oxide during RIE and adjusting the sidewall oxide thickness prevents stress from the bottom corners to appear

. increasing the temperature of field and capping oxidation prevents screw dislocation from the top corners to occur.

- polysilicon planararisation.

Done with a sacrificial resist layer, the etch back is stopped just before reaching the nitride, not to damage it. Then, an appropriate oxidation of the remaining polysilicon is performed. A wet etch clears the so-formed oxide and a flat surface is reached. One can notice that the expansion of polysilicon in the trench during that last oxidation can be used to fill the possible voids in between the two polysilicon strips.

A conventional locos technique is employed to cap the polysilicon in the trenches and to realize the field oxide. After the Si_3N_4 stripping, the deep N+ region is phosphorus implanted and driven. Next, extrinsic base is BF2 implanted. Subsequently a plasma deposited SiO_2 reinforces the base oxide, then emitter and contacts are opened. A 200nm thick amorphous silicon layer is deposited at 560°C under low pressure with a cleaning procedure which leads to an interfacial oxide having a minimum thickness. Then the intrinsic base and the emitter are implanted (boron and arsenic respectively). After silicon etch, a 10mn 960°C thermal treatment gives a 70nm deep emitter in the monosilicon, with a 100nm wide base.

A 80nm thick PtSi in formed in the monosili-

con contacts and at the top of the polysilicon in order to realize one interconnect level inside the cells. A 500nm thick plasma deposited layer of SiO_2-P_2O_5 achieves the isolation between the PtSi and the first metal. The metallic interconnect system is realized with two layers of AlSiTi. The pitches are 4µm and 6µm for the first and the second metal layer respectively. A combination of polyimide and SiN with tapered VIA procedure ensures the isolation between metallic layers.

3. EXPERIMENTAL RESULTS

A SEM cross-sectional view of a NPN transistor is shown in figure 2. The transistor area is 33µm². The emitter area is 0.75 x 1.5µm². The defects generation at the corners of the trenches has been monitored with the wright etch delineation. After optimisation of the filling conditions and the subsequent thermal treatments no longer defects have been detected as shown in the figure 3.



FIGURE 2 - SEM cross-sectional view of the NPN transistor

The various transistor structures designed allow an accurate determination of the most important electrical parameters.

For example Table 1 summarizes transistor characteristics with different sizes.

Figure 4 shows forward current gain of T6 transistor, which is designed with 0.7µm rules.

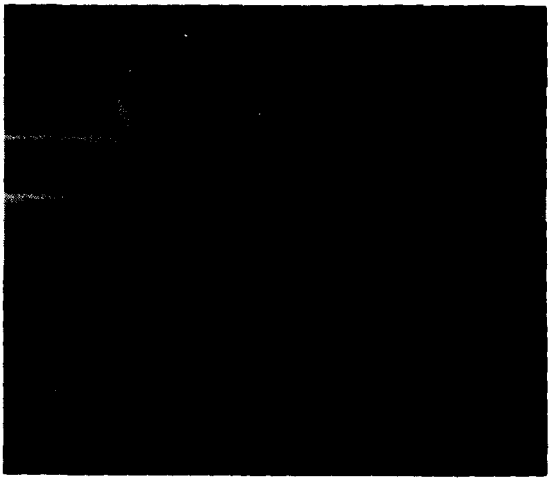


FIGURE 3 - SEM cross-sectional view of the trenches after Wright etch.

TRANSISTOR	T6	TRT2	TRT3
Emitter Area (µm²)	0.75x1.5 = 1.125	1x2.5=2.5	1x5.5=5.5
Base Area (µm²)	7x1.5=10.5	11.75x2.5 = 29.4	13x5.5=71.5
Collector Area (µm²)	11x1.5 = 16.5	15.75x2.5 = 39.5	11.25x5.5 = 61.9
BF	160	80	110
IKF (mA)	1.6	3.8	7.5
RE (Ω)	70	40	20
Rbb' (Ω)	1700	250	100
C_{BE} (fF)	4.2	16.2	27.2
C_{BC} (fF)	7.1	19.5	25.2
C_{CS} (fF)		8	18
Cut-off Frequency (GHz)		18	18
BVEBO (V)	6	6.5	6.5
BVCBO (V)	12	13	13
BVCEO (V)	10	11	11

TABLE 1 - Main transistor characteristics.

One notices the very good uniformity of hfe from 1µA to 100µA.

The cut off frequency Ft, (measured on transistor TRT3 which is large enough to eliminate parasitic effects [6] : bonding pads capacitances, rbb'...), is 18 GHz at VCE = 5V and IC = 600µA (Figure 5).

The comparison between measured and simula-

ted delay times of ring oscillators, made of differential ECL inverters (21 stages) allows us to validate the above parameters (maximal deviation about 5%).

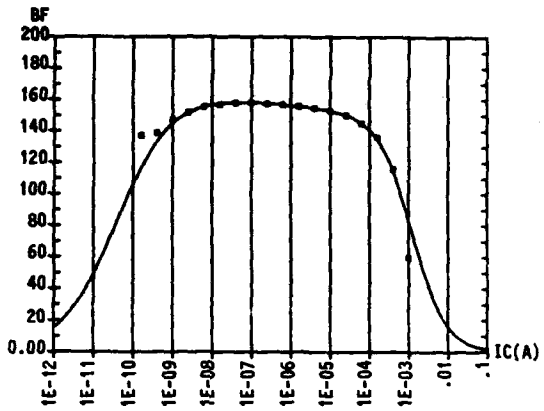


FIGURE 4 - Forward current gain versus collector current at VBC = 0 (NPN T6 Em = 0.75 x 1.5µm²)

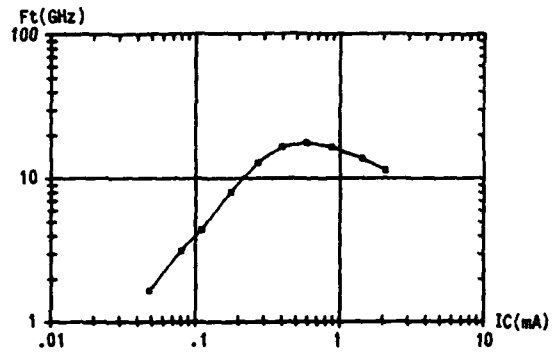


FIGURE 5 - Ft versus IC at VCE = 5V (NPN TRT3 Em = 1 x 5.5µm²)

Table 2 gives the results of ring oscillators ROV2 and ROV3 at the supply voltage VEE = 2V.

Ring Oscillator	Transistor type	Emitter Size (µm)	Propagation Delay time (ps)	Current/ stage (mA)
ROV2	TRT2	1 x 2.5	115	0.4
ROV3	TRT3	1 x 5.5	95	1

TABLE 2 - ECL ring oscillators, VEE = 2V Internal logic swing ΔV = 350 mV

Fig. 6 shows the propagation delay/stage and the internal logic swing, versus the current/ stage for ROV3 (VEE = 2V).

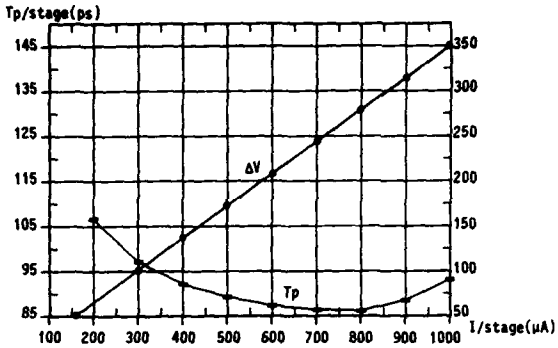


FIGURE 6 - Tp/stage and logic swing ΔV at VEE = 2V (Ring oscillator ROV3).

4. CONCLUSION

A new bipolar process devoted to applications requiring sub 100ps gate delays and very high packing density has been set up. The isolation is achieved with dielectric filled deep trenches. An emphasis has been put on the reduction of the defects associated with the trench filling. The polysilicon emitter has proved to be effective to achieve high performances with good yield. A 18 GHz cut off frequency has been measured. The minimum propagation delay time evaluated in ECL ring oscillator is 88 ps.

This work is supported by the E.E.C program ESPRIT under Contract n° 243 : "Submicron Bipolar Technology".

REFERENCES

[1] M. VORA et al, IEDM Tech. Dig. (1985) p. 34

[2] M. SUZUKI et al, IEEE J. Sol. Stat. Circuits Vol SC22, N°1, (1987) p. 41

[3] H.K. PARK et al, IEEE El. Dev. letters, Vol ED27, N°12 (1986)

[4] Tamaki et al, Jap. J. Appl. Phys. 21 (1981)

[5] D.D. TANG et al, IEEE J. Sol. Stat. circuits Vol SC17, N°5, (1982) p. 925

[6] H.M. REIN, IEEE J. solid state Electronics Vol 26, N°1, (1983) p. 75

Session B2.3

SOI Workshop III

Chairman: D. Mc Caughan

Tuesday, September 15, 1987

SOI STRUCTURES BY ION IMPLANTATION AND ANNEALING IN A TEMPERATURE GRADIENT

G. K. CELLER

AT&T Bell Laboratories Murray Hill, NJ 07974, USA

Si-on-insulator (SOI) technology is essential for space and military applications where immunity to ionizing radiation is required. It will also lead to commercial, high speed CMOS circuits. This paper reviews formation of SOI structures by implantation of high doses of oxygen, followed by heat treatments at 1405°C in a lamp furnace. The result is an SiO_2 layer, with atomically abrupt interfaces, buried under a film of a device quality single crystalline Si. Since the SOI technology and the 3-D integration place new requirements on control of dopants in the wafers, trapping and drift of arsenic introduced into the SiO_2 is also discussed. It is demonstrated that As is immobile in the oxide under isothermal conditions up to the melting temperature of silicon. When a temperature gradient is applied, the arsenic is swept toward the heat source until it reaches the SiO_2/Si interface and is released into Si. The difference between dopant transport with and without temperature gradients adds flexibility to processing of multilayer structures.

1. INTRODUCTION

Si-on-insulator (SOI) technology provides better immunity to ionizing radiation, simpler design, increased density of components, higher switching speed, and elimination of latchup [1-3]. The benefits of SOI are best realized in CMOS technology. This is illustrated in Fig. 1, where a conventional bulk CMOS inverter is drawn next to one formed on a Si island that is dielectrically isolated from the substrate. The layout of both inverters assumes the same minimum feature size. The SOI structure is smaller because the deep diffusion tubs are eliminated. The number of lithographic masks required to make the SOI circuits is also reduced.

1.1. Scope of the paper.

In this paper, formation of high quality Si-on-SiO_2 films by oxygen implantation and very high temperature annealing is discussed [4,5]. The direct synthesis of a buried oxide has emerged in the last few years as the SOI method that is the most likely to be implemented in radiation-hard MOS electronics. Commercial VLSI applications are expected to follow some years later.

In the second part of the paper, some unique properties of arsenic in SiO_2 are discussed [6]. Since As is an important n-type dopant of Si, the flexibility of trapping and controllably releasing the As from the oxide into Si is important for the SOI technology [7].

The best SOI structures by oxygen implantation require heat treatments at the highest possible temperatures. Transport of As in the buried oxide requires a high temperature and a temperature gradient. Both requirements are satisfied in a special experimental configuration described below.

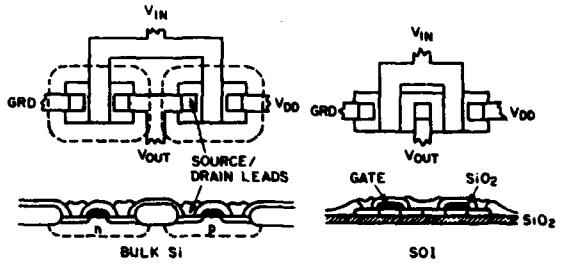


FIG. 1. A comparison of CMOS inverters in bulk Si and in Si-on-SiO_2.

2. HEATING AT HIGH TEMPERATURE AND IN A TEMPERATURE GRADIENT

Annealing at 1405°C was performed in a lamp furnace, with the samples suspended between a bank of high intensity lamps and a water cooled base [8]. The back side of each sample was held exactly at the melt temperature of silicon, T_M=1412°C, as indicated schematically in Fig. 2,

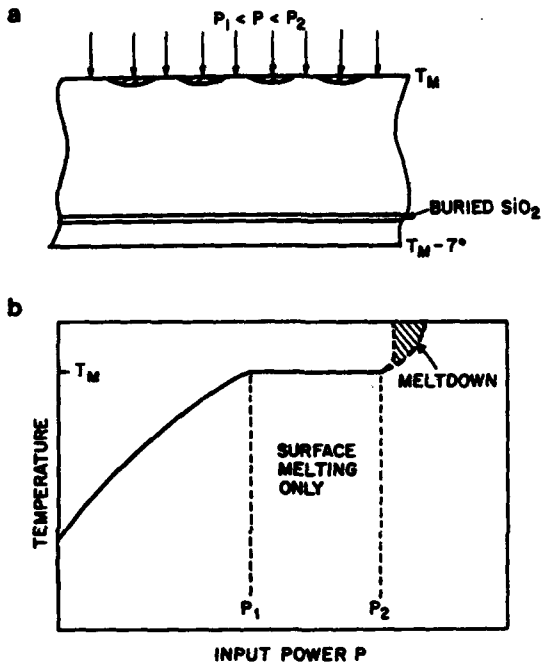


FIG. 2. The heating arrangement with an intrinsic temperature stability. (a) A cross section of a radiatively heated sample, with one surface partially molten and the implanted surface at a slightly lower temperature. (b) Three temperature regions, representing conventional heating below T_M, constant temperature regime, and deep melting.

and the implanted side was 7°C cooler because of the temperature gradient imposed by the heating configuration [9]. This gradient is a function of temperature T, emissivity ϵ of the surface not exposed to light, and thermal conductivity κ,

$$\nabla T = \epsilon \sigma \frac{T^4}{\kappa}$$

where σ is the Stefan-Boltzmann constant. At 1400°C the thermal conductivities of Si and SiO_2 are ~0.2 W/(cm K) and ~0.02 W/(cm K), respectively, resulting in temperature gradients of ~0.014°C/μm in Si and ~0.14°C/μm in SiO_2 when $\epsilon \approx 0.6$ is assumed for the Si surface. The temperature stability is based on the change in optical properties of silicon upon melting and ensures that the surface exposed to the photon flux will remain at T_M without any risk of overmelting [10].

3. SOI WAFERS BY OXYGEN IMPLANTATION

Formation of oxides and nitrides by ion implantation has been investigated since the late 1960s [5]. The first major effort to prepare SOI wafers by oxygen implantation was initiated at NTT by Izumi, who coined the acronym SIMOX, for Separation by IMplanted OXygen [11]. A high dose of oxygen ions is implanted deep into silicon in such a way as to leave a surface film of single crystalline Si above a layer that eventually becomes SiO_2. This requires a dose of above 10^{18} cm^{-2} oxygen ions at an energy of at least 100 keV, but preferably 150-200 keV, to form silicon dioxide films buried under 1000 to 3000 Å of crystalline silicon [12,13]. The implantation is carried out at an elevated temperature, in 450 to 600°C range, to preserve crystallinity of the surface layer of silicon.

As the oxygen dose is increased, the shape of oxygen distribution changes from a skewed Gaussian profile typical for conventional ion implantation to a flat-topped one. This change is caused by oxygen saturation at a stoichiometric value of oxygen in SiO_2, indicating that SiO_2 is formed already during implantation. Oxygen exceeding the stoichiometric concentration of 4.4 x 10^{22} cm^{-3} diffuses rapidly toward the wings of the distribution, where it oxidizes more Si. The critical dose for forming the SiO_2 during 200 keV implantation is 1.4 x 10^{18} cm^{-2}.

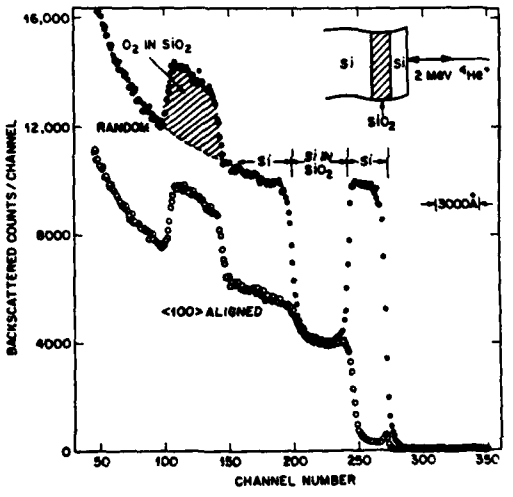


FIG. 3. RBS random and <100> aligned channeling spectra for a sample implanted with molecular oxygen at 400 keV to a dose of 1.8 x 10^{18} cm^{-2} and annealed for 30 min at 1405°C.

After implantation, an additional heat treatment is required to repair damage caused in the silicon overlayer and to diffuse more oxygen toward the SiO_2 where it is incorporated in the oxide film. The final SOI microstructure is strongly coupled with the annealing conditions [9,14,15]. Annealing above 1300°C is needed to obtain sharp Si/SiO_2 interfaces. The heat treatment at 1405°C, near the melting point of Si, is particularly effective in chemical segregation of implanted oxygen into the buried SiO_2 and in recovering crystallinity of the Si film [4,9].

The results shown below have been obtained by implanting Si (100) wafers at ~550°C with O_2^+ at 400 keV to a dose of 1.8×10^{18} cm^{-2}, using the Heavy Ion Accelerator at the University of Surrey, or by implanting 1.7×10^{18} of O^+ at 150 keV using the Eaton NV-200 high current implanter. After implantation the wafers were coated with an oxide film, of 0.2-0.5μm thickness, to protect Si films during annealing from oxidation or pitting.

Annealing at temperatures up to 1250°C was carried out in a conventional, resistively heated tube furnace. Heat treatments at 1405°C were carried out in the lamp furnace described above.

Rutherford backscattering (RBS) and channeling spectra after a 30 min heat treatment at 1405°C, shown in Fig. 3, indicate that the entire silicon layer is a high quality single crystal, with the minimum yield $\chi_{min} = 3.3\%$. The random spectrum shows that the oxygen has redistributed completely to form a SOI structure with sharp and well defined boundaries between silicon and oxide [16].

RBS and channeling are not sensitive to dislocations. Cross-sectional and plan-view transmission electron microscopy (TEM) are necessary to examine the details of the microstructure. In Fig. 4, a series of sections after 2h

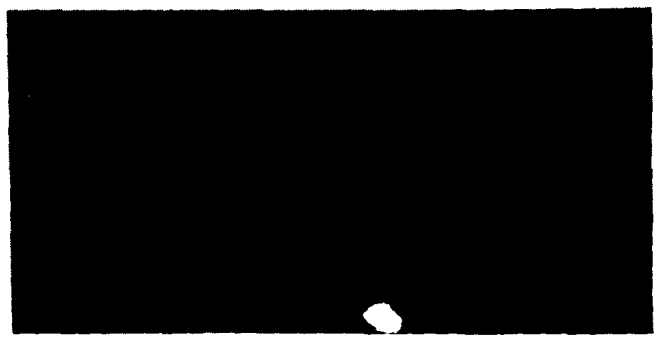


FIG. 4. TEM cross sections of three samples, which were implanted with the same oxygen dose [17].

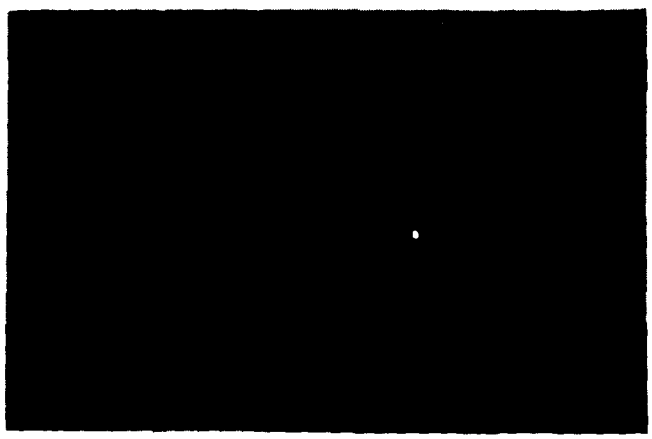


FIG. 5. TEM cross section of a sample implanted at a lower energy of 150 keV.

at 1150, 2h at 1250, and 0.5h at 1405°C are shown [17]. After annealing at 1150°C the stoichiometric oxide layer is ~40% thinner than would be expected from the implanted dose. Both oxide interfaces are very irregular. The SiO_2/substrate interface consists of a series of oxide platelets separated by defective silicon. The silicon film above the oxide has two regions, the top 700Å is crystalline with the threading dislocations density of 3×10^9 cm^{-2}, and the the remaining 3000Å is highly defective and contains amorphous oxide polyhedra. After 1250°C the density of threading dislocations is reduced to 3×10^8 cm^{-2}, but large oxide precipitates remain in the Si film. Smaller precipitates have been dissolved, contributing to the growth of a few large precipitates and increasing the thickness of the planar buried oxide layer. Both Si/SiO_2 interfaces are smoother than at the lower temperatures, but there are still considerable undulations. Platelets near the lower interface have been replaced by a row of large silicon precipitates trapped inside the SiO_2 and aligned in a row about 300Å away from the interface.

After a treatment at 1405°C no precipitates can be detected in silicon on either side of the oxide. The thickness of Si and oxide films is 2800 and 3700Å, respectively. Both oxide interfaces are smooth, but some faceted crystallites of Si still remain embedded in the oxide near the back interface. This can be seen more clearly in a cross sectional view of another sample, shown in Fig. 5. This sample has been implanted at 150 keV, with the temperature held near 580°C. Because of the lower implant energy and the 750Å screen oxide used during implantation, the Si film over the oxide is only about half the thickness of that in the previous Figure. Other features are very similar; sharp interfaces, the absence of oxide precipitates in Si and the row of Si precipitates near the back oxide interface. Threading dislocations are the main defects left in the Si film.

Many groups are currently building and characterizing device structures in SIMOX films, from individual transistors and ring oscillators to 64K static random access memories [3,18]. The majority of these devices was made in films without abrupt interfaces, which necessitated deposition of epitaxial Si to increase the thickness of the single crystalline films. Since wafers processed above 1300°C have a thicker layer of precipitate-free silicon, epitaxial deposition may no longer be necessary. CMOS transistors in such wafers have excellent electrical parameters [19].

Recently, van Ommen et al. have reduced the density of dislocations below $10^5 cm^{-2}$ by modifying implantation conditions for the maximum dose uniformity [20]. If this result can be extended to practical ion dose rates, the quality of SIMOX films will be further improved.

4. PHASE SEPARATION AND DRIFT OF As IN SiO_2 IN A TEMPERATURE GRADIENT

Diffusion of arsenic in the SiO_2 is a complex phenomenon [21-24]. It is shown here that transport of arsenic in the oxide can be controlled by parameters other than the temperature alone. Arsenic implanted into the SiO_2 can remain immobile at temperatures exceeding 1400°C, or can be swept rapidly in one direction, depending on the heating conditions. The new degree of control over diffusion of As may be useful in building SOI structures.

Kinsbron et al. [25,26] observed that during a complete oxidation of an As-implanted polysilicon film on SiO_2, small arsenic-rich precipitates had formed at the initial location of the polysilicon/SiO_2 interface. The arsenic was pushed to this interface by the advancing oxidation front.

More recently we have observed that a large dose of As implanted directly into SiO_2 precipitates inside the oxide into a separate, As-rich phase, when heated in an oxygen-free ambient [6,27]. The spherical As-rich droplets propagate in a temperature gradient by the mechanism of thermomigration. By controlling not only the temperature but also the ∇T we can move the entire As-rich zone toward the heat source or hold it essentially immobile.

4.1. Experimental Conditions

Arsenic doses of 3×10^{14} to 3×10^{16} cm^{-2} were implanted at 100 keV into SiO_2 grown in steam at 1050°C on (100) Si wafers. For some experiments, additional SiO_2 was deposited over the thermal oxide by low pressure CVD to bury the implanted arsenic deeper inside the oxide. A 1.5μm layer of polycrystalline silicon was deposited on the oxide at 620°C. This Si film is necessary to increase surface emissivity over that of the oxide, since this controls the ∇T. The silicon film also acts as a diffusion barrier for any oxygen left in the furnace chamber and as a sink for mobile oxygen present in the SiO_2.

Samples were heated from one side only in the lamp furnace. The sample temperature was above 1200°C, and in most experiments near 1405°C, with the $\nabla T \cong 0.14$°C/μm in the SiO_2.

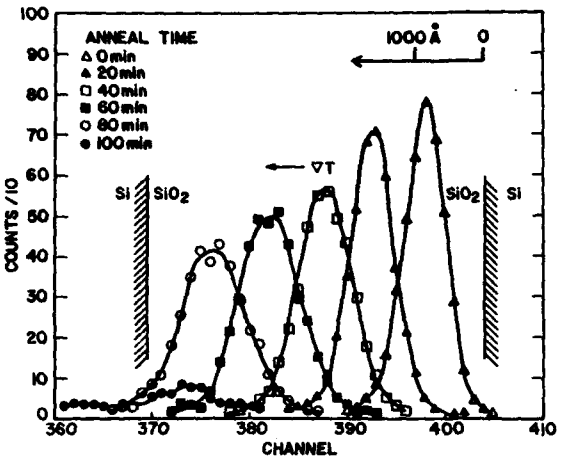


FIG. 6. Distribution of arsenic in the SiO_2 after different heating periods in the temperature gradient. The origin of the distance scale is at the oxide surface after removal of the polycrystalline Si coating.

4.2. Arsenic Drift Data

RBS results for a series of heat treatments at 1405°C of a sample with 10^{16} As implanted into 0.4μm of SiO_2 and coated with 1.5μm of polycrystalline Si are shown in Fig. 6. The peaks represent arsenic before the first heat treatment and after a series of 20 min heating cycles at 1405°C. The center of the As distribution moves at a linear rate of 2300Å/hour, with little loss of mass, and with broadening that is small compared to the peak shift. By measuring the area under the peaks we have determined that only 4% of As is lost from the drifting zone during the first 60 minutes. After 100 min the entire peak reaches the interface and outdiffuses into silicon.

Cross-sectional TEM micrographs from the same series of samples show that As has segregated into spherical inclusions of 10 to 30 nm diameter and that these inclusions move in the SiO_2 towards the heat source, as shown in Fig. 7. The As-rich spheres and the SiO_2 are amorphous; there is no indication of impurity induced crystallization.

The drift velocity is approximately the same for all implanted doses of arsenic above 3×10^{15} cm^{-2}. For lower As doses, there is no segregation or drift, and RBS spectra indicate conventional diffusion.

Results for samples with thicker oxide, and in particular for samples where the As was buried deep inside the SiO_2 are somewhat different. In these samples there is either little or no drift during an incubation period, from 30 to 120 min depending on the placement of arsenic. After the incubation period, the drift is at a constant rate which is about the same as that in the thin oxide samples from Figures 6 and 7. TEM micrographs show that in these samples the spherical As-rich inclusions form slowly, and only after they reach a critical radius of about 50Å, the drift begins.

4.3 Discussion of As Transport Phenomena

In the absence of free oxygen, arsenic in SiO_2 segregates into spherical inclusions, provided that the initial As concentration is sufficient (above 1 at. %). The distance between the arsenic-rich band and the Si/SiO_2 interfaces determines the incubation time for the growth of these



FIG. 7. TEM micrographs of As in the SiO_2. (a) Before high temperature treatment; (b) after 20 min, and (c) after 40 min at 1405°C The temperature gradient is 0.14°C/μm with the heat supplied from the bottom in this figure. Distortion of the upper Si/SiO_2 interface is caused by grain growth in polycrystalline silicon.

inclusions to the critical size needed for drifting in the temperature gradient. We interpret this last dependence as related to the rate of removal of oxygen (possibly present in the form of OH or H_2O) from the vicinity of As atoms. At high temperatures, this oxygen diffuses through the SiO_2 until it reacts with Si at one of the two Si/SiO_2 interfaces. The resulting oxygen concentration gradients cause the most rapid depletion of oxygen near the interfaces and the slowest drop in concentration at the center of the SiO_2.

If a temperature gradient is present in addition to the elevated temperature, the droplets drift with a constant velocity toward the heat source, until they reach the Si/SiO_2 interface and conventional diffusion of As in Si takes over.

The drift is caused by thermomigration, a process which typically involves liquid droplets moving through the solid by dissolution of the solid at the hotter side of the droplet, diffusion of the matrix atoms through the liquid, and their redeposition on the cooler side [28]. The rate of droplet migration is limited either by diffusion in the liquid or by kinetics of dissolution and redeposition. The constant velocity of the As-rich droplets, independent of the droplet radius above a 50Å threshold, is consistent with the drift governed by volume diffusion in the liquid [6].

4.4. Applications of As drift

Drift of arsenic in the SiO_2 can be utilized to form buried conducting layers near the isolation oxide. In the example discussed here, 50μm thick Si films were recrystallized from the melt over 4μm thick oxide [29]. The high voltage device structure required that n^+ region be formed in the recrystallized layer, adjacent to the buried oxide. Since the Si film has to be melted, it can only be doped after resolidification, but the film thickness precludes ion implantation into the recrystallized layer. The solution is to implant the arsenic into the SiO_2 before deposition of the polycrystalline Si overlayer [7]. It remains immobile inside the oxide through the entire high temperature processing sequence. After recrystallization, a heat treatment in a temperature gradient sweeps the As distribution out of the SiO_2 and into silicon. Concentration of As in Si, for two directions of the ∇T is shown in Fig. 8. A similar approach, but with shorter heat treatments in a ∇T, could be used in thin film SOI structures. In the case of multiple active layers of Si on SiO_2, the n^+ dopant held inside the oxide films would be released near the end of the processing sequence. Directed energy heating could also be employed for a spatially localized arsenic transport in a large ∇T.

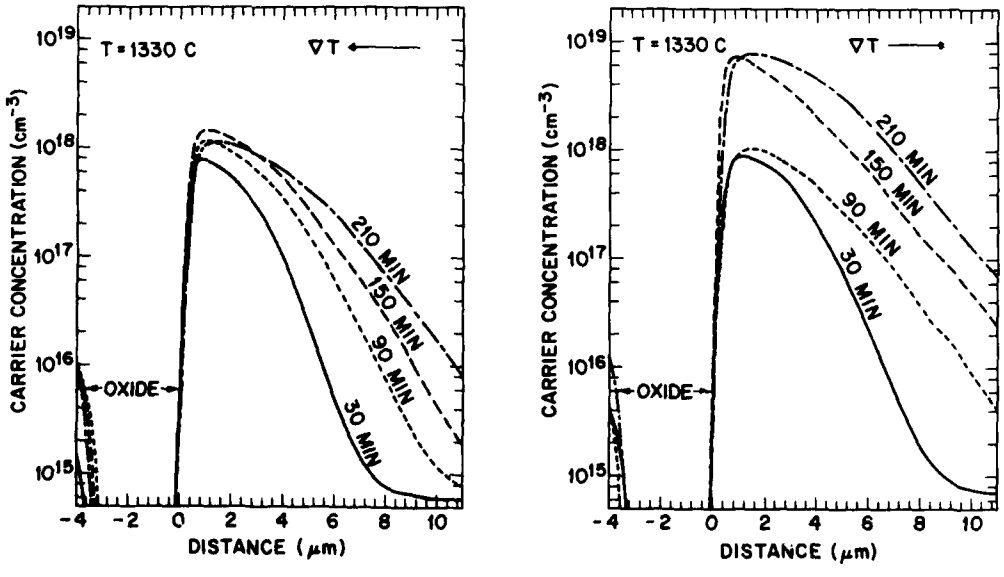


FIG. 8. Formation of a buried n^+ region in Si adjacent to the isolation oxide. The amount of As that outdiffuses from the SiO_2 depends strongly on the direction of the temperature gradient.

5. SUMMARY

High temperature annealing has a profound effect on the microstructure of SOI films formed by high dose oxygen implantation. Oxygen precipitates dissolve rapidly near 1400°C and oxygen outdiffuses from the Si film to the interface with the buried oxide, where it forms more SiO_2. This eliminates the broad band of oxygen-rich defective Si, typical of lower temperature anneals. Because of better quality and homogeneity of the Si film, an additional deposition of epitaxial Si is no longer necessary.

The discovery of the drift of As in the SiO_2, and the new understanding of the As diffusion phenomena in the oxide, provide a new tool for doping of Si devices. This should be particularly useful in SOI structures, which inherently contain a buried oxide layer.

ACKNOWLEDGMENTS

The author thanks J. L. Batstone, L. Pfeiffer, T. T. Sheng, L. E. Trimble, and K. W. West of AT&T Bell Labs, P. L. F. Hemment and K. J. Reeson of the University of Surrey, and C. D. Marsh of Oxford University for their contributions.

REFERENCES

[1] *Single-Crystal Silicon on Non-Single-Crystal Insulators*, edited by G. W. Cullen, J. Cryst. Growth **63**, 429-582 (1983).

[2] Mater. Res. Soc. Symp. Proc. **53**, A. Chiang. M. W. Geis, and L. Pfeiffer, eds., (Materials Research Soc., 1986).

[3] G. K. Celler, Proc. of the 1st Int. Symp. on ULSI, S. Broydo and C. M. Osburn, eds., (The Electrochem. Soc. 1987).

[4] G. K. Celler, J. L. Batstone, K. W. West, P. L. F. Hemment, and K. J. Reeson, Proc. European MRS, Vol. XII, G. G. Bentini, E. Fogarassy, and A. Golanski, eds., (Les editions de physique, 1986), p. 95.

[5] P. L. F. Hemment, Mater. Res. Soc. Symp. Proc. **53**, 207 (1986).

[6] G. K. Celler, L. E. Trimble, K. W. West, L. Pfeiffer, and T. T. Sheng, Mater. Res. Soc. Symp. Proc. **92**, S. R. Wilson, R. Powell and D. E. Davies, eds. (Materials Research Soc., 1987).

[7] G. K. Celler and L. E. Trimble, Proc. Symp. on High Voltage and Smart Power Devices, P. W. Shackle, ed., (The Electrochem. Soc. 1987).

[8] G. K. Celler, CRC Critical Reviews in Solid State an Mater. Sciences, **12**, 193 (1984).

[9] G. K. Celler, P. L. F. Hemment, K. W. West, and J. M. Gibson, Appl. Phys. Lett. **48**, 532 (1986).

[10] G. K. Celler, McD. Robinson, L. E. Trimble, and D. J. Lischner, Appl. Phys. Lett. **43**, 868 (1984).

[11] K. Izumi, M. Doken, and H. Ariyoshi, Electron Lett. **14**, 593 (1978).

[12] K. Izumi, Y. Omura, and S. Nakashima, Mat. Res. Soc. Symp. Proc. **23**, 443 (1984).

[13] H. W. Lam and R. Pinizzotto, J. Cryst. Growth **63**, 554 (1983).

[14] C. Jaussaud, J. Stoemenos, J. Margail, M. Dupuy, B. Blanchard, and M. Bruel, Appl. Phys. Lett. **46**, 1064 (1985).

[15] B. -Y. Mao, P.-H. Chang, H. W. Lam, B. W. Shen and J. A. Keenan, Mater. Res. Soc. Symp. Proc. **53**, 251 (1986)

[16] G. K. Celler. P. L. F. Hemment, K. W. West, and J. M. Gibson, Mat. Res. Soc. Symp. Proc. **53**, 227 (1986).

[17] C. D. Marsh, G. R. Booker, K. J. Reeson, P. L. F. Hemment, R. J. Chater, J. A. Kilner, J. A. Alderman, and G. K. Celler, Proc. European MRS, Vol. XII, G. G. Bentini, E. Fogarassy, and A. Golanski, eds., (Les editions de physique, 1986), p. 137.

[18] Abstracts of 1986 IEEE SOS/SOI Technology Workshop (Captiva Island, Oct. 1986).

[19] J. R. Davis, K. Reeson, P. L. F. Hemment, and C. D. Marsh, IEEE Device Letters (1987).

[20] A. H. van Ommen, H. J. Ligthart, J. Politiek, and M. P. A. Viegers, Mater. Res. Soc. Symp. Proc. **93**, U. Gibson, P. P. Pronko, A. E. White, eds.,

[21] J. Wong and M. Ghezzo, J. Electrochem. Soc. **119**, 1413 (1972).

[22] Y. Wada and D. A. Antoniadis, J. Electrochem Soc. **128**, 1317 (1981).

[23] R. Singh, M. Maier, H. Krautle, D. R. Young, and P. Balk, J. Electrochem. Soc. **131**, 2645 (1984).

[24] A. H. van Ommen, J. Appl. Phys. **56**, 2708 (1984).

[25] E. Kinsbron and W. T. Lynch, IEDM Tech. Digest, p. 674 (IEEE 1983).

[26] E. Kinsbron, S. P. Murarka, T. T. Sheng, and W. T. Lynch, J. Electrochem. Soc. **130**, 1555 (1983).

[27] G. K. Celler, L. E. Trimble, K. W. West, L. Pfeiffer, and T. T. Sheng, Appl. Phys. Lett. **50**, 664 (1987).

[28] H. E. Cline and T. R. Anthony, J. Appl. Phys. **48**, 2196 (1977).

[29] G. K. Celler, McD. Robinson, L. E. Trimble, and D. J. Lischner, J. Electrochem. Soc. **132**, 211 (1985).

OXYGEN IMPLANTATION FOR SOI DEVICE TECHNOLOGIES

Kevin YALLUP*, An De VEIRMAN†, Luc DUPAS‡, Kristin De MEYER‡+

* Analog Devices, c/o IMEC Kapeldreef 75, B-3030 Leuven
† Universiteit Antwerpen, RUCA, Groenborgerlaan 171, B-2020 Antwerpen
‡ IMEC Kapeldreef 75, B-3030 Leuven

In this work the formation of silicon on insulator (SOI) substrates by oxygen implantation (SIMOX) and annealing followed by an epitaxial deposition step will be described. Physical results for the material will be presented and then the corresponding electrical parameters will be discussed and compared with those obtained from bulk wafers processed simultaneously. From the results described we hope to indicate the areas in which problems are likely to occur when using such substrates in an IC process.

1 SAMPLE PREPARATION

Standard CZ n-type, (100), 5 ohm-cm wafers implanted with oxygen ions to a dose of $2x10^{18}cm^{-2}$ at an energy of 150 keV and a temperature of 600 °C have been obtained from a commercial source. The substrates were then capped with 100 nm of CVD oxide in order to protect the surface and annealed at a temperature of 1250 °C for times of 0, 2, 4 and 8 hours in a nitrogen ambient. After annealing, the capping layer was stripped in buffer HF. One specimen of each annealing time was prepared for TEM examination as described elsewhere [1]. A second set of substrates were given the 8 hour anneal after which a modified version of a standard process was used to deposit a 5 μm epitaxial layer of lightly n-doped silicon. Only the longest annealing time was used for this part of the work in order to obtain the best quality starting material for the growth step. One specimen was also prepared for TEM examination and the remaining wafers were processed to form devices.

A simple p-well CMOS process was used to manufacture test structures on both the SIMOX substrates described above and bulk control wafers which had received an identical epitaxial deposition. No threshold adjust implants were employed, the n-channel threshold was set by the doping level of the well and the p-channel value was set by that of the substrate. The gate oxide thickness was 100 nm. In order to further simplify processing, field isolation between devices was provided by a simple guard-banding scheme in conjunction with a deposited thick oxide.

2 STRUCTURAL RESULTS

Figure 1a shows a composite cross sectional TEM micrograph of the 0,2 and 4 hour annealed specimens. It can be seen that the most striking changes in material properties of the buried oxide and silicon layers take place in the first 2 hours of the annealing cycle. The very large number of small silicon dioxide precipitates seen after implantation are replaced by a much smaller number of larger precipitates and the oxide/silicon interfaces become much more abrupt. In addition, threading dislocations are formed in the surface semiconductor layer. After the 4 hour anneal, the oxide interfaces become somewhat smoother and the oxide precipitates are completely eliminated from the silicon overlayer. This is confirmed by the plan view TEM micrographs of the 2 and 4 hour heat treatments shown in Figures 1b and 1c. In these images the oxide precipitates are labelled P and crystalline silicon carbide precipitates C. After 8 hours, very little further change in the structure of the SIMOX substrate is observed.

In contrast to the oxide precipitates, the threading dislocation density does not appear to change significantly during the course of the heat treatments studied here. One possible reason for the apparent stability of these structures is the fact

+ Research Associate of the Belgian National Fund for Scientific Research, Professor at the K.U. Leuven



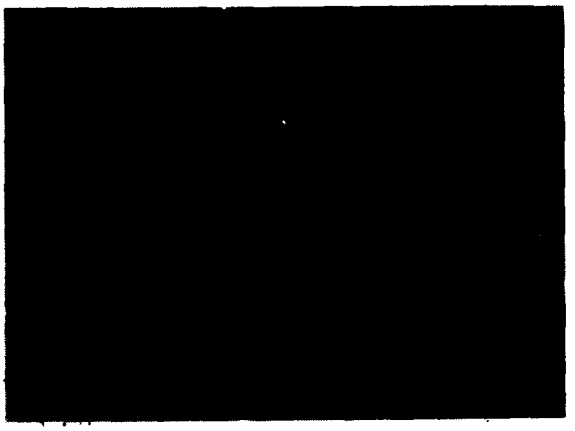




FIGURE 1

1a) Cross-sectional HVEM image of the oxygen implanted structures after annealing for 0, 2 and 4 hours at 1250 C.

1b), 1c) plan view TEM of silicon overlayer after 2 hours (1b) and 4 hours (1c) anneal at 1250 C.



FIGURE 2

Cross-sectional HVEM image of the epitaxial layer illustrating the decrease of dislocation density towards the surface.

that they traverse the silicon overlayer from one interface to the other. A more complete description of the formation and evolution of the defect structures described above can be found in [1]. In general, our findings are very similar to those that have been observed in studies carried out by other workers e.g. [2,3,4] at both similar and higher temperatures.

Structural examination of a wafer annealed for 8 hours followed by an epitaxial deposition in cross-sectional TEM (Figure 2) shows that the threading dislocations present after implantation and annealing propagate into the deposited layer. It can be seen that the density of such defects appears to decrease from the bottom to the top. A surface density of approximately $7x10^{7}cm^{-2}$ has been obtained by secco etching and SEM examination. This value should be compared with an estimated post-anneal surface dislocation density of $3x10^{9}cm^{-2}$ which has been measured on plan view TEM micrographs such as Figure 1. Clearly some elimination of dislocations is taking place

during the course of the epitaxial deposition. More detailed TEM studies have shown dislocations apparently merging and annihilating each other but at present the mechanism for this phenomenon is not clear. Further work is required to clarify the exact processes that take place during the deposition of epitaxial layers on SIMOX substrates.

3 ELECTRICAL RESULTS

Gate oxide to silicon interface state density has been estimated by comparing high and low frequency (quasi-static) C-V measurements made on large area capacitors. The results are shown in Table 1. Clearly the surface state density at the oxide/silicon interface is not significantly increased by the presence of large numbers of defects such as those seen in Figure 2. Furthermore, measurements on long channel transistors (Table 1) show that neither the channel mobility nor the threshold voltage of such devices are significantly affected by the oxygen implantation. Such a result is to be expected for

		SIMOX	BULK
D_{it} $ev^{-1}cm^{-2}$	p-well	2.4 10^{10}	1.8 10^{10}
μ [cm^2/Vs]	n-chan	730 ± 7	734 ± 8
	p-chan	318 ± 3	321 ± 4
Vt [V]	n-chan	1.16 ± 0.01	1.25 ± 0.01
	p-chan	-1.10 ± 0.03	-1.10 ± 0.02

Table 1: Electrical measurements on large square transistors.

the SIMOX substrates in this study as the epitaxial layers employed are thick enough that the buried oxide will not influence the performance of surface devices such as MOS transistors.

On the other hand, one of the bulk properties that can be expected to be modified by the presence of large numbers of defects is lifetime. C-t measurements of large area gate oxide capacitors were carried out and Zerbst analysis has been used to derive a value for the carrier lifetime. Table 2 compares values of this parameter for the two substrate types. It can be seen that the SIMOX substrates do indeed have a shorter minority carrier lifetime than bulk substrates.

It is well known that the carrier lifetime can also have an influence on the leakage of junctions through the generation component of the current. Table 2 compares the diode leak-

		SIMOX	BULK
τ_{gen} μs	p-well	3	10
Irev Acm^{-2}	n^+	$7x10^{-5}$	$1x10^{-7}$
	p^+	$2x10^{-7}$	$2x10^{-7}$
Ideality	n^+	1.09	1.05
	p^+	1.09	1.05

Table 2: Electrical measurements on large area diodes and capacitors. Reverse bias leakage was measured at 15 V.

age of bulk and oxygen implanted wafers at 15V reverse bias. It can be seen that the SIMOX n^+ to p-well junctions have a poorer performance than the bulk equivalents whereas the p^+ to n-substrate devices give approximately the same leakage current in both cases. More detailed examination of the reverse-bias characteristics show that the n^+ diodes in the oxygen implanted substrates have very soft characteristics with a slightly lower breakdown voltage than the bulk equivalents. On the other hand, when the forward bias ideality factor is examined in Table 2 it can be seen that both the n^+ and the p^+ diodes are similarly degraded in the oxygen implanted substrate. Other workers [5] have also reported similar observations. Such an apparently contradictory result can be understood if it is remembered that the forward and reverse biased diode currents are made up from two terms, due to generation and diffusion respectively, weighted by different factors in the two conduction regimes. Thus, in the case of reverse bias leakage, the experimental results suggest that the generation current is only significant for the n^+ diodes, whereas in the forward bias regime it appears to be an important component of the total conduction current of both types of device.

Another problem that could occur in junctions formed in highly dislocated material is the formation of spikes by the enhanced diffusion of dopant along the core of such defects. Differences in diffusion behaviour of the n and p dopant species could thus offer an alternative explanation for the differences in diode reverse bias leakage. However, the possibility of spiking has been largely ruled out in the present study by examination of the temperature dependence of the reverse bias leakage behaviour. If this problem were present, it would be expected that the leakage would be determined by the electric field at the spike tip and thus be relatively independent of temperature. Five measurements of the temperature dependence of the reverse bias current at 5V were carried out on each substrate

type. In all cases the results showed an Arrhenius type behaviour with an activation energy of approximately 1 eV. Such a value is not consistent with leakage currents due to junction spikes.

4 DISCUSSION

We have shown that the technology of high current oxygen implantation combined with a suitable annealing cycle and epitaxial deposition step is capable of producing SOI substrates where the only significant defects are threading dislocations. These structures are present in the top silicon region after the annealing process and propagate through the epitaxial layer during deposition. However, we have shown that the dislocation density apparently decreases from the back of this layer to the front. The density at the surface is $7x10^{7}cm^{-2}$.

It is known that dislocations generally only have a weak electrical activity in themselves. However, it is likely that they can act as preferential sites for the formation of more detrimental defects [6]. In particular, heavy metals (e.g. Cu, Fe) can be gettered by such structures to form electrically active precipitates.

In the present study, we indeed observe a reduction in lifetime of the SIMOX material relative to bulk controls such as might be expected in the presence of decorated dislocations. We might reasonably anticipate that other important device performance parameters influenced by generation lifetime will be altered in SIMOX material and in fact junctions fabricated in oxygen implanted material show a somewhat degraded performance when compared to those made on bulk wafers. The results we have obtained are similar to those reported in [6] where the change in diode parameters was attributed to decorated stacking faults. On the other hand, if parameters that are not directly affected by the generation lifetime are examined, such as channel mobility or threshold voltage, then no difference is seen between bulk and SIMOX substrates.

Although the degradation of lifetime in SIMOX substrates that we have reported is undesireable, it is probably not large enough to result in an unacceptable loss of MOS device performance. However, the value of this parameter measured in this work is rather low in the oxygen implanted wafers when compared to that expected for a process in which the fabrication schedule has been optimised to improve the lifetime. Thus it is likely that such substrates will not be suitable for the manufacture of high performance bipolar or power devices unless some improvement in material quality can be achieved.

5 CONCLUSIONS

We have shown that oxygen implanted substrates can be processed in a relatively conventional manner to form starting material of acceptable quality. It is possible to process such wafers using a simple CMOS process to form devices with an acceptable performance. On the other hand, the material does not appear to be suitable for the fabrication of bipolar or power devices unless some substantial improvement can be made in the quality of the as-implanted and annealed susbstrates.

ACKNOWLEDGEMENTS

We should like to thank Dr. H. Maes and Dr. J. Van Landuyt for their assistance with this work.

References

[1] De Veirman A., Yallup K., Van Landuyt J., Maes H.E., Amelinckx S., Proc. 5th Oxford Confr. Microsc. Semicond. Mat., to be published in Inst. Phys. Confr. Series 1987

[2] Hemment, P.L.F., European Mat. Res. Soc. Proc. 1985 p475

[3] Bruel M., Jaussaud C., Margail J., and Stoemenos J., European Mat. Res. Soc. Proc. 1986 p105

[4] Marsh C.D., Hutchinson J.L., Booker G.R., Reeson K.J., Hemment P.L.F., Celler G.K., Proc. 5th Oxford Confr. Microsc. Semicond. Mat., to be published in Inst. Phys. Confr. Series 1987

[5] Krull W.A., Rouse G.V., Cherne R.D., Buller J.F., 1986 IEEE SOS/SOI Workshop

[6] Dishman J.M., Haszko S.E., Marcus R.B., Murarka S.P., Sheng T.T., J. App. Phys. 50(4) 1979 p2689

SMALL GEOMETRY SOI/CMOS DEVICES ON SIMOX SUBSTRATES

J R Davis, K Reeson* and P L F Hemment*

British Telecom Research Labs, Martlesham Heath, Ipswich, UK.
*University of Surrey, Guildford, Surrey, UK.

Small geometry CMOS transistors have been made in SOI wafers produced by high dose oxygen implantation. By performing the implantation at high energy (200 keV) and annealing the wafers at 1300°C, the thickness and quality of the resulting silicon film is such that the expensive and difficult to control step of epitaxial growth is not needed. The lack of any major crystallographic defects (other than threading dislocations) results in the absence of any anomolous lateral diffusion of the source/drain dopants, allowing 1 micron gates to be used without excessive channel shortening.

1. INTRODUCTION

Considerable progress has been made in recent years towards producing the silicon on insulator (SOI) layers by oxygen implantation. The primary application for these layers is as a substrate for high-performance CMOS circuits.

In many cases in the literature [1], the CMOS transistors have been fabricated in an epitaxial layer grown on top of the thin silicon film isolated by the oxygen implantation. The advantage of this technique is that the active areas of the devices avoid any crystalline damage left in the silicon film after implantation and annealing. The disadvantages are that the silicon islands are thicker than optimum and also that the epitaxial growth step is expensive and difficult to control. In this paper we show that by using a high energy (200 keV) implant, followed by annealing at 1300°C, the quality of the silicon layer and its interfaces is sufficient to support high-performance transistor action without the need for epitaxy.

2. METHOD

Standard 3" (100) 20 ohm.cm p-type wafers were implanted with 1.8×10^{18} O^+ ions/cm^2 at 200 keV. Only the central 2.5 x 2.5 cm^2 area was implanted and the wafers were maintained at temperatures of around 500-520°C by beam heating. No screen oxide was used. The wafers were then annealed at 1300°C for 2 or 20 hours in a nitrogen ambient and using a protective cap of deposited oxide. CMOS transistors were fabricated in these wafers, without epitaxy, using a conventional polysilicon gate process with LOCOS isolation. The total reduction in the thickness of silicon islands due to the growth of various oxides was 34 nm. Boron implantation was performed with a range of doses into the channels of both the p- and n-channel transistors to adjust the threshold voltages and to control conduction along the back channel. The gate oxide thickness was 27 nm.

3. MATERIALS

TEM cross-sections [2] of wafers similar to those used for device fabrication showed that the thickness of the buried oxide was approximately 400 nm, in good agreement with the value of 410 nm predicted by assuming that the total implanted dose is converted to stoichiometric silicon dioxide. The top silicon film was approximately 290 nm thick. Wafers annealed for 2 hours (at 1300°C) showed an abrupt interface between the buried oxide and the silicon film. There was, however, a row of small (20-40 nm) oxide precipitates in the silicon film about 100 nm from the buried

interface. These precipitates disappeared in the specimen annealed for 20 hours, in agreement with the results of Margail et al [3]. Threading dislocations, running through the total thickness of the silicon film, were found for both annealing times. The measured densities of these dislocations were 1×10^9 cm^{-2} after 2 hours and 4×10^8 cm^{-2} after 20 hours but it is not certain if this reduction is attributable to the increased annealing or to small variations in the implantation conditions.

Although the TEM cross-sections and RBS studies showed that the longer annealing time produced an improvement in the crystal structure, SIMS profiling [4] indicated that it was accompanied by a peak of high oxygen concentration very near the top surface of the silicon. The cause of this effect is not known and is being investigated further. Probably as a result of this high oxygen level, the electrical properties of the top interface were slightly degraded by the longer anneal. The results that follow refer to the 2 hour anneal unless otherwise stated.

4. ELECTRICAL RESULTS

4.1 General

Low field inversion mobilities were extracted from plots of reciprocal gain versus drawn channel length. Average values of 580 cm^2/V.s for electrons and 205 cm^2/V.s for holes were obtained from n- and p-channel transistors with optimum boron implant doses. These values are within a few percent of those obtained on bulk transistors with similar doping levels and are better than those obtained on earlier wafers annealed at 1150°C or 1200°C.

Electron mobilities were also measured at the back interface by using the silicon substrate as a gate. For the wafer annealed for 20 hours, values in the range 150-450 cm^2/V.s were obtained, far higher than for earlier devices [5]. The highest values of back-channel mobility corresponded to the highest implantation temperatures. This indicates that better regrowth of the back interface and dissolution of the oxide precipitates occurs if amorphisation of the silicon near the buried oxide is avoided [6].

Drain junction leakages measured with the transistors in the 'off' state and with $|V_D|$ = 1V were typically below 5×10^{-14} A/micron channel width for both transistor types. Even for p-channel transistors in which the channel was buried by using a boron counter-doping implant, the leakage did not exceed 10^{-13} A/micron. It is probable that these low values are obtained in the presence of the large number of dislocations because they do not intersect the junction.

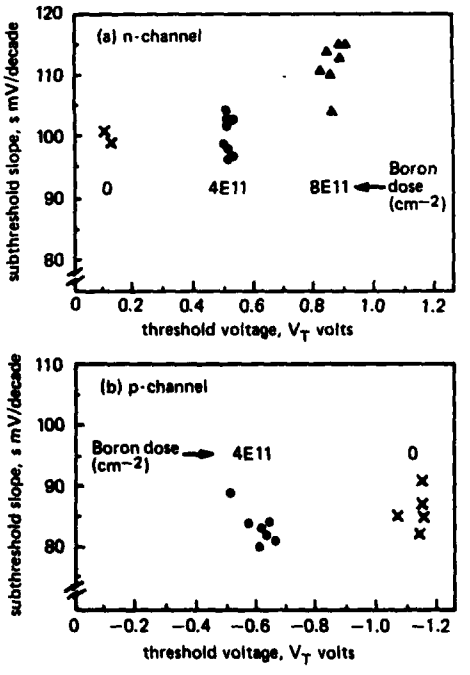


FIGURE 1

Correlation between the threshold voltage and subthreshold slope of individual SOI transistors for a variety of shallow boron implant doses. (a) n-channel, (b) p-channel.

The correlation of the subthreshold slope s (s = dV_G/d (log(I_D)) in mV/decade) and the threshold voltage is shown in Figure 1 for several channel implants. Values of less than 100 mV/decade are achieved with the single exception of the n-channel transistor with the highest implant dose. These values are in good agreement with those predicted by a modified version of MINIMOS [7] assuming an ideal silicon film. Earlier transistors in films annealed at lower temperatures had significantly worse subthreshold slopes and could be modelled by assuming the presence of bulk states in the silicon.

The subthreshold characteristics of the n-channel transistors shown in Fig. 1 were measured with a bias applied to the back gate (V_{BG} = -10V) in order to prevent conduction along the edges of the transistor at the

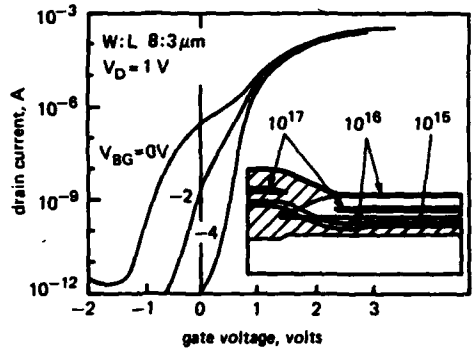


FIGURE 2

Subthreshold characteristics of an n-channel transistor as a function of the back-gate voltage. Inset: calculated channel doping profile (cm^{-3}) at the island edge.

interface with the buried oxide. In Fig 2 an anomolous low-level component of subthreshold current is seen to be removed by accumulating the back interface of the film with a negative voltage. The characteristics of edgeless (closed geometry) transistors were independent of this bias. The cause of the extra current at the device edges is the change in doping profile caused by the bird's beak, as shown in the inset of Fig 2. This reduction in effective doping can be overcome by simple changes in the dopant implantation schedule, thus rendering back-gate bias unnecessary. Note that the edge-current path in this case with LOCOS-isolated islands is at the back interface, in contrast to the island edge effects reported previously for mesa-isolated devices [8]. In this latter case the current flows along the top edges of the islands where the gate fields from the top and sides of the island overlap and cause a local reduction in threshold voltage.

4.2 Short Channel Effects

Output characteristics of short channel transistors (1.0 micron gates) are shown in Fig 3. Both transistor types show good current drive capability and relatively small

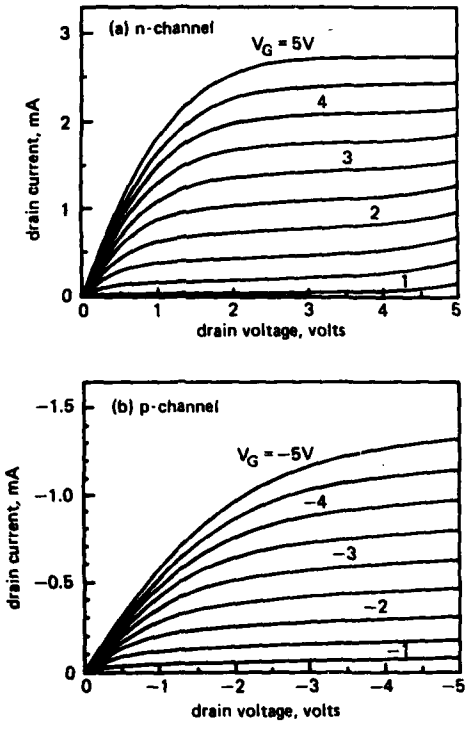


FIGURE 3

Output characteristics for short channel transistors in SOI wafers annealed at 1300°C for 2 hours. Gate length = 1 micron, channel width = 8 microns. (a) n-channel, (b) p-channel.

short channel effects. No 'kink' is seen in the n-channel characteristics due to the fact that with the relatively thin silicon film the island is completely depleted and there is no neutral region to be charged by hot carrier effects [9].

The values of ΔL, the difference between the nominal gate length and the electrical channel length, were measured as 0.1 and 0.2 microns for the n- and p-channel transistors respectively. Corresponding values for the back-channel transistors were -0.1 and 0 microns. These results indicate that there is no significant-anomolous lateral diffusion of the source and drain dopant along the back interface. The fact that the effective channel lengths at the back interface are larger than those at the front suggest that the source and drain regions have their normal semi-circular shape. This is in contrast to observations in silicon on sapphire where fast diffusion along twin boundaries leads to greater lateral spread at the back interface than at the surface [10].

The reduction in n-channel threshold voltage as the gate length is decreased from 10 to 1.0 microns is only about 50 mV and remains well controlled for all values of channel doping. The p-channel transistors with a surface channel (no boron implant) also show well-controlled thresholds for gate lengths down to 1.0 microns. The buried channel transistors however, show the effects of punch-through for gate lengths below 1.5 microns. Two possible solutions of this problem are the use of a deep phosphorus implant or the use of thinner silicon films.

5. CONCLUSIONS

These results demonstrate that the combination of performing the oxygen implantation at 200 keV and annealing at very high temperatures results in an SOI substrate which is capable of supporting high-performance CMOS transistors without the need for epitaxy. The processing sequence used, although by no means optimised, shows that viable short channel transistors can be produced with very simple processing.

ACKNOWLEDGEMENTS

The authors would like to thank Dr S Cristoloveanu (ENSERG, Grenoble) for suggesting the technique of measuring the effective channel length at the back interface, R Chater and Dr J Kilner for SIMS profiles, and C Marsh for TEM cross sections. This work was partially funded by the UK Alvey programme and the Science and Engineering Research Council (SERC). J Davis acknowledge the permission of the Director of Research, British Telecom, to publish this paper.

REFERENCES

[1] Izumi, K., Omura, Y. and Nakashima, S., Electron. Lett., Vol. 22, No. 15, p. 775, 1986.

[2] Davis, J.R., Reeson, K., Hemment, P.L.F. and Marsh, C.D., IEEE Electron Dev. Lett., Vol. 8, No. 7, 1987.

[3] Margail, J., Stoemenos, J., Jaussaud, C., Depuy, M., Martin, P., Blancard, B. and Bruel, M., presented at the Mater. Res. Soc. Meeting, Strasbourg, France, May 1985.

[4] Chater, R. and Kilner, J., Imperial College, London, U.K., unpublished.

[5] Cristoloveanu, S., Lee, J.H., Pumfrey, J., Davis, J.R., Arrowsmith, R.P. and Hemment, P.L.F, J. Appl. Phys., Vol. 60, No. 9, p. 3199, 1986.

[6] Tuppen, C.G., Taylor, M.R., Hemment, P.L.F. and Arrowsmith, R.P., Appl. Phys. Lett., Vol. 45, No. 1, p. 57, 1984.

[7] Armstrong, G.A. and Davis, J.R., in the proceedings of NASECODE IV, The Fifth International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuits, Trinity College, Dublin, Ireland, June 1987.

[8] Foster, D.J., Electron. Lett., Vol. 19, No. 17, p. 684, 1983.

[9] Colinge, J.-P., Electron. Lett., Vol. 22, No. 4, p. 187, 1986.

[10] Yallup, N.E., Cowan, K.J. and Godfrey, D.J., Appl. Phys. Lett., Vol. 48, No. 11, p. 70, 1986.

CARRIER GENERATION AND TRAPPING PROPERTIES IN SIMOX STRUCTURES

Tarek Elewa, Hisham Haddara and Sorin Cristoloveanu

Laboratoire de Physique des Composants à Semiconducteurs,
ENSERG, 23 rue des Martyrs, 38031 Grenoble Cedex, France.

New methods and results related to the characterization of silicon on insulator material formed by deep oxygen implantation (SIMOX) are presented. An original and exact analysis allows us to determine both the minority carrier lifetime and the surface recombination velocity by monitoring the transient drain current of depletion mode SIMOX transistors pulsed in deep depletion. In addition, the dynamic transconductance technique, initially proposed for enhancement devices, is adapted to characterize interface properties of depletion mode transistors.

1. INTRODUCTION

The characterization of interface and volume properties of silicon on insulator (SOI) material is difficult using conventional capacitance methods. This is due to : (1) the complex multi-interface nature of SOI structures, (2) the absence of a film contact, (3) the increased influence of parasitic capacitances [1] and (4) the very small gate area of VLSI devices. In this paper, we present new results related to the characterization of SIMOX materials fabricated under different processing conditions. These results concern several important parameters, namely : the minority carrier lifetime τ, the surface recombination velocity s_o as well as the density of interface traps. Such a characterization has been possible by means of two original methods which we present, for the first time, in the next two sections.

2. MINORITY CARRIER LIFETIME AND SURFACE RECOMBINATION VELOCITY

The minority carrier lifetime τ is a very important parameter in characterizing the quality of silicon films. The application of the standard deep depletion capacitance technique [2] is not straightforward due to the strong influence of parasitic capacitances and series resistance [1]. This difficulty has been avoided by monitoring the transient drain current in a depletion mode transistor (DMT) while being pulsed in deep depletion [3]. However, the extraction of τ was carried out using an approximate analysis which neglects the influence of surface recombination. Here, we present an exact analysis, similar to that of Zerbst [2], to determine both τ and s_o.

From the equation of charge conservation in an MOS system, the sum of the inversion charge Q_{inv} and the interface trap charge Q_{it} is :

$$(Q_{inv}+Q_{it}) = C_{ox}(V_G+V_{FB}) - \frac{qN_dC_{ox}}{2\epsilon_s} W^2 - qN_dW \qquad (1)$$

with C_{ox} : the gate oxide capacitance, V_G : the gate voltage, V_{FB} : the flat band voltage, N_d : the film doping, W : the depletion layer width and ϵ_s : the silicon permitivity. The depletion layer width W can be expressed in terms of the drain current I_d as follows :

$$W = W_f - \frac{L}{\sigma Z V_D} I_d = W_f - KI_d \qquad (2)$$

where W_f is the silicon film width, L : the channel length, Z : the channel width, σ : the film conductivity, V_D : the drain voltage and $K = L/(\sigma Z V_D)$. By differentiating (1) with respect to time we get :

$$-\frac{N_dC_{ox}}{2n_i\epsilon_s} \frac{d}{dt}\left[\left(W_f+\frac{\epsilon_s}{C_{ox}}\right) - KI_d(t)\right]^2 = \frac{K[I_d(\infty)-I_d(t)]}{\tau} + s_o \qquad (3)$$

where $I_d(\infty)$ is the final steady state current. The right hand side (R.H.S.) of (3) represents the generation in the bulk as well as under the gate. By plotting the left hand side (L.H.S.) of (3) versus $[I_d(\infty) - I_d(t)]$, we obtain a straight line from which τ and s_o can be extracted using the slope and the intercept respectively.

Measurements were performed on SIMOX samples fabricated by implanting a high oxygen dose (1.8×10^{18} O^+/cm^2) at 200 keV with an implantation temperature of 500 °C. Both low temperature (1200 °C) and high temperature (1300-1400 °C) anneals have been performed. A semiconductor parameter analyzer (HP 4145B) was used to monitor the transient drain current as a function of time. Fig. 1 shows the transient drain current for an edgeless N^+NN^+ transistor with $N_d = 10^{17}$ cm^{-3} and annealed at 1400 °C. The values of τ and s_0, obtained from the linear part of the transient shown in Fig.2, were 1.6 µs and 0.2 cm/s respectively. The high value obtained for τ gives evidence of the good quality of the silicon film after annealing at high temperature. In contrast, much smaller lifetime values (10-100 ns; according to processing conditions) were obtained for silicon films annealed at low temperatures which confirms their inferior quality. The departure from linearity at the beginning of the transient (Fig.2) is attributed to a non constant surface recombination velocity. The low value obtained for s_0, from the intercept, corresponds to the surface recombination velocity under the gate when the surface is strongly inverted. Measurements were also performed on conventional transistors (i.e. with edges) ; the results indicate the strong influence of the sidewalls generation component on the slope of the straight line. This suggests a new definition of an "effective" measured lifetime. More details will be given elsewhere.

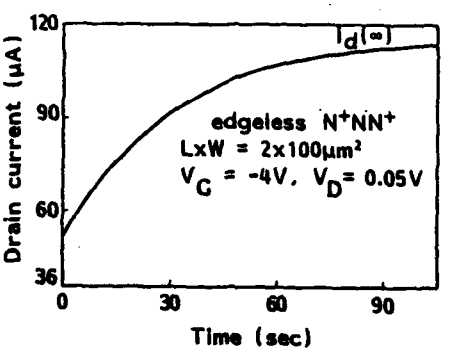


FIGURE 1

Drain current as a function of time for a depletion mode transistor pulsed in deep-depletion (HTA-SIMOX, V_D = 50 mV, V_G = -4 V)

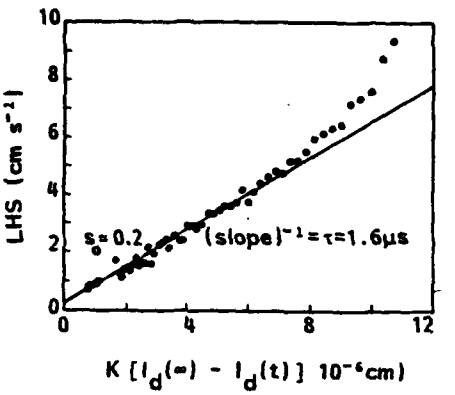


FIGURE 2

Extraction of the minority carrier lifetime τ and the surface recombination velocity s_0 from the slope and the intercept respectively (see eqn.3).

3. INTERFACE TRAP PROPERTIES

Recently, a new technique has been proposed for the determination of interface trap properties from the dynamic transconductance of enhancement MOSFETs made on bulk-Si [4]. In this section, the dynamic transconductance technique is adapted to depletion mode transistors and applied to characterize transistors fabricated on SIMOX material.

In an N^+NN^+ transistor which is not totally depleted and operates in the ohmic region (i.e. $V_D \rightarrow 0$), the drain current is given by :

$$I_d = \mu \frac{Z}{L} (Q_{dmax} - Q_d) V_D \qquad (4)$$

with $Q_{dmax} = qN_dW_f$. The transconductance $g_m = dI_d/dV_G$ is, therefore given by :

$$g_m = \mu \frac{Z}{L} V_D \frac{dQ_d}{d\Psi_s} \cdot \frac{d\Psi_s}{dV_G} \qquad (5)$$

where Ψ_s is the surface potential.

In the case of harmonic small signal operation, $d\Psi_s/dV_G$ is obtained by differentiating the instantaneous charge conservation equation, in the depletion regime, with respect to time :

$$\frac{d\Psi_s}{dV_G} = \frac{C_{ox}}{C_{ox} + C_d + Y_{it}/(j\omega)} \qquad (6)$$

where ω is the angular frequency of the applied gate signal, C_d : the depletion layer capacitance and Y_{it} : the equivalent parallel interface trap admittance ($Y_{it} = G_p + j\omega C_p$ [5]). By substituting (6) into (5) and replacing dQ_d/dV_s by C_d, we get :

$$g_m(\omega) = K_1 \frac{C_d C_{ox}}{C_d + C_{ox} + Y_{it}/(j\omega)} \qquad (7)$$

with $K_1 = \mu(Z/L)V_D$. Similar to enhancement MOSFETs, G_p/ω is simply obtained from the imaginary part of $1/g_m(\omega)$:

$$\frac{G_p}{\omega} = K_1 C_d C_{ox} \, \mathrm{Im}[1/g_m(\omega)] \qquad (8)$$

The depletion capacitance C_d can be known from the real part of $1/g_m(\omega)$ at high frequency. In the case of enhancement MOSFETs operating in weak inversion, G_p/ω is given by [4]:

$$\frac{G_p}{\omega} = \frac{C_{ox} I_d}{kT/q} \, \mathrm{Im}[1/g_m(\omega)] \qquad (9)$$

Dynamic transconductance measurements were performed on enhancement and depletion mode SIMOX transistors. Fig.3 shows a plot of $\mathrm{Im}[1/g_m(\omega)]$ versus frequency for an enhancement MOSFET implanted at 480 °C and annealed at low temperature (1200 °C). The double peak behaviour is due to the presence of both interface and bulk traps. The interface trap peak shifts towards higher frequencies with increasing V_G while the other peak does not change neither in position nor in magnitude and is, therefore, attributed to bulk traps. Fig.4 shows G_p/ω versus frequency for an another enhancement MOSFET implanted at 515 °C and annealed at 1400 °C. The bulk traps peak dissappears and an average interface trap density of 4×10^{10} $eV^{-1}cm^{-2}$ is deduced from the maxima of the peaks.

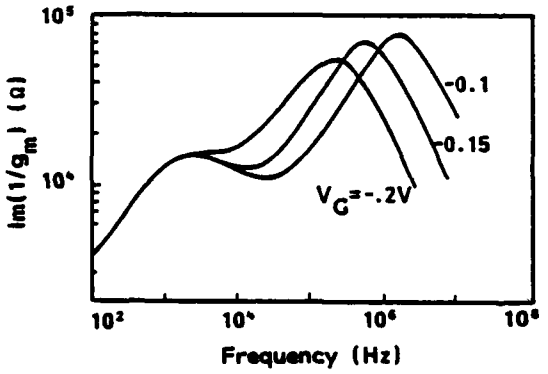


FIGURE 3

Variation of the imaginary part of $1/g_m(\omega)$ as a function of frequency, with V_G as a parameter in weak inversion, for an enhancement mode transistor on SIMOX material implanted and annealed at low temperature.

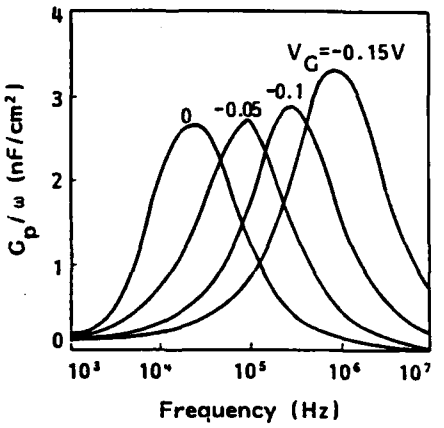


FIGURE 4

G_p/ω versus frequency for an enhancement mode transistor with an implantation temperature of 515 °C and annealing temperature of 1400 °C.

The qualities of the front and back interfaces, in a depletion MOSFET are compared in Fig.5. In case of the back interface,

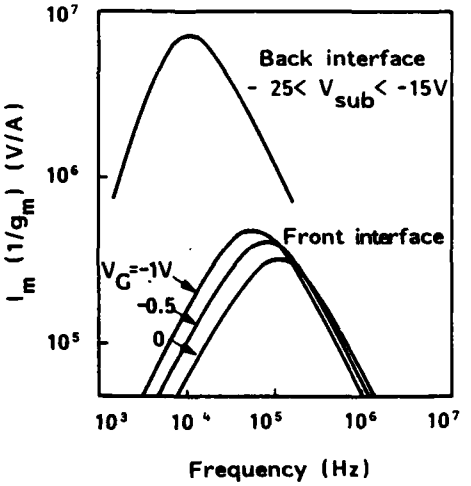


FIGURE 5

Experimental plot of $\mathrm{Im}[1/g_m(\omega)]$ versus frequency, in depletion, for a depletion mode transistor. The implantation temperature was 515 °C while the annealing temperature was 1400 °C.

bulk traps (present in the depletion region above the interface) dominate. This is evident since the obtained peak does not change neither in position nor in magnitude with varying V_G. On the contrary, only interface trap peaks are obtained for the front interface with an average interface trap density of 10^{11} $eV^{-1}cm^{-2}$. Due to the complexity of the equivalent circuit including bulk traps, the relationship between the bulk trap admittance and $g_m(\omega)$ is not straightforward. Consequently, quantitative information about them is more difficult to obtain than for interface traps. A seperate study will be devoted to this problem.

4. CONCLUSIONS

We have presented new methods for the characterization of interface and volume recombination properties in SIMOX material. The dynamic transconductance method is adapted to depletion mode transistors and applied, for the first time, to determine the interface trap properties in both enhancement and depletion type devices on SIMOX material. The minority carrier lifetime as well as the surface recombination velocity are determined using a depletion mode MOSFET pulsed in deep depletion. Finally, the effect of implantation and annealing temperatures on the quality of silicon films was investigated.

REFERENCES

[1] J.H.Lee, and S.Cristoloveanu, IEEE electron Dev. Lett., EDL-7 (1986) 537.

[2] M.Zerbst, Z. Angew. Phys., 22 (1966) 30.

[3] D.P.Vu and J.C.Pfister, Appl. Phys. Lett., 9 (1985) 951.

[4] H.Haddara and G.Ghibaudo, Europhys. Conf. Abstracts, 10G, 47, ESSDERC'86, Cambridge, UK (1986).

[5] E.Nicollian and G.Goetzberger, The Bell Syst. J. 46 (1967) 1055.

SMALL GEOMETRY NMOS TRANSISTORS IN SILICON-ON-SAPPHIRE USING RAPID ANNEALED SOURCE/DRAINS AND IMPROVED CRYSTALLINE QUALITY SILICON FILMS

M Field, N E B Cowern* and D J Godfrey

GEC Research Limited, Hirst Research Centre, East Lane, Wembley, Middlesex HA9 7PP

*Now at Phillips Research Laboratories, Eindhoven, Netherlands

1. INTRODUCTION

The advantages of silicon-on-insulator (SOI) over bulk silicon technology for CMOS devices are well documented [1] and include greater packing densities, increased radiation hardness, immunity from latch-up and lower parasitic capacitances. At present silicon-on-sapphire (SOS) is the only readily available SOI substrate. SOS material has differences with respect to single crystal silicon substrates which must be taken into account during processing as they ultimately affect device characteristics. One of the most important differences is the defect density profile. In SOS the lattice mismatch between the silicon and the sapphire results in the presence of crystalline defects in the silicon, the concentration of which decreases with distance from the interface as the mismatch is accomodated. The defects are mostly stacking faults, micro-twins and dislocations in the epitaxial silicon layer. It has been found previously [2] that dopant diffusion in SOS can be increased over bulk silicon behaviour by the presence of defects.

For the formation of small geometry MOSFET's (gate lengths < 1 μm), it is necessary to limit the source/drain diffusion not only in the lateral direction under the gate electrode, but also in depth in order to limit deleterious short channel effects. Experimental and modelling studies of dopant redistribution in as-grown SOS have shown that the residual defects result in enhanced diffusion compared to equivalent bulk silicon behaviour. For example, Figure 1 illustrates a part of a cross-section of an MOS device in which the gate electrode has acted as an implant mask for a self-aligned arsenic source/drain implant on the left. The contours show the arsenic concentration after a 900°C 30 minute furnace anneal predicted by the SOS diffusion model. The bottom of the diagram is the silicon/sapphire interface.

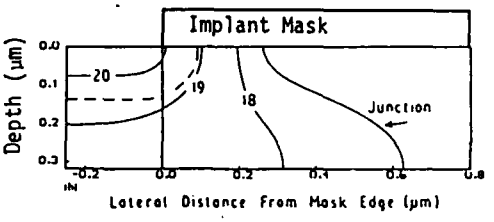


Fig 1: 40 keV 5×10^{15} cm^{-2} arsenic 900°C 30 min anneal. Solid lines give carrier profile in as grown SOS and dashed line gives junction contour for equivalent bulk silicon diffusion

It can be seen that there is;

(i) An increased sideways diffusion of the source/drain region into the channel under the gate region, in comparison with the bulk silicon equivalent.

(ii) A lateral extension of the source/drain region near the silicon/sapphire interface.

The first effect increases the parasitic gate capacitance, while it is expected that the second contributes to the off-state leakage of SOS devices, by creating a leakage path along

the silicon/sapphire interface. This diffusion behaviour may be restricted by either removing the crystalline defects present in the silicon layer or by using short duration, high temperature heat treatments for annealing out the source/drain implantation damage.

In the current work, NMOS devices have been fabricated on both 0.3 μm and 0.6 μm thick as-received SOS wafers and improved crystalline quality SOS obtained using the double solid phase epitaxy and growth (DSPEG) technique [3]. A mesa island etched NMOS process with a gate oxide thickness of 300 A was used to produce various size transistors from 20 μm to 0.75 μm length (20 μm width).

The anneal and activation of the source/drain regions was performed using either a 900°C 30 minute furnace treatment in a dry nitrogen ambient or rapid optical heating to (nominally) 1100°C for five seconds. Table 1 gives a summary of the major long channel device characteristics. These are averages of about fifty devices in each case, the errors are approximately two standard deviations of the distribution (i.e. absolute error).

In particular, it can be seen that the DSPEG material results in the highest channel mobilities (indicative of the removal of crystalline defects) and, as anticipated, that the sideways diffusion (the difference between the physical and electrical channel lengths) is reduced for both DSPEG and rapid annealed cases.

To investigate the effect of the lateral diffusion of dopant at the silicon/sapphire interface, the off-state leakage currents of the NMOS devices were measured. Each device had its gate electrode held at 1 Volt below the measured threshold voltage for that device whilst the drain leakage current was measured for a source-drain potential difference of 5 Volts. In this way the off-state leakage current of a device could be plotted against the electrical channel length of that device.

Figure 2,3 and 4 show the results obtained for 0.3 μm as-received SOS, 0.3 μm DSPEG material and 0.3 μm as-received SOS using the rapid anneal. For the first case the leakage rises rapidly for electrical channel lengths below approximately 1.3 μm, whereas for the other cases the electrical channel lengths below which leakage rises is less than 1 μm. The larger scatter in Figure 4 is thought to be due to the greater variations in temperature of the rapid optical anneal between wafers. By considering the electrical channel length, the differences in sideways diffusion near the $Si-SiO_2$ interface between samples are largely accounted for and the observed behaviour may be attributed to the differences in arsenic diffusion at the silicon/sapphire interface.

In conclusion, it has been shown that either using the DSPEG technique or reducing the thermal cycling of the source/drain anneal, good working NMOS SOS transistors with sub-micron electrical channel lengths can be obtained. The dopant profiles after diffusion are being investigated and will be presented along with detailed electrical characterisation.

Table 1:

SOS substrate	Threshold voltage (V)	Mobility (cm^2/Vsec)	Sideways diffusion (μm)
		Furnace Anneal	
0.6 μm as received	0.70 ± 0.02	336 ± 20	0.12 ± 0.02
0.3 μm as received	0.59 ± 0.02	328 ± 10	0.12 ± 0.02
0.3 μm DSPEG	0.52 ± 0.02	416 ± 15	0.06 ± 0.02
		Rapid Optical Anneal	
0.6 μm as received	0.58 ± 0.10	336 ± 20	0.08 ± 0.02
0.3 μm as received	0.64 ± 0.03	269 ± 20	0.07 ± 0.03
0.3 μm DSPEG	0.47 ± 0.05	386 ± 15	0.07 ± 0.03

ACKNOWLEDGEMENTS

This work was partly supported by the Alvey Directorate and the rapid optical annealing was performed at STL.

REFERENCES

[1] Partridge S.L., IEDM 1986, Tedchnical Digest, Section 16.6, 428-430

[2] Cowern N.E.B., Yallup K.J. and and Godfrey D.J., Appl. Phys. Lett., 48 (11), 1986, 704-706

[3] Peters T.B. and Dineen C. IEEE SOS/SOI Technology Workshop, Florida, 1986,

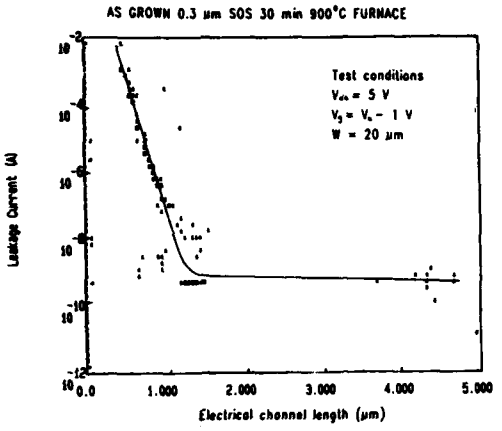


Figure 2

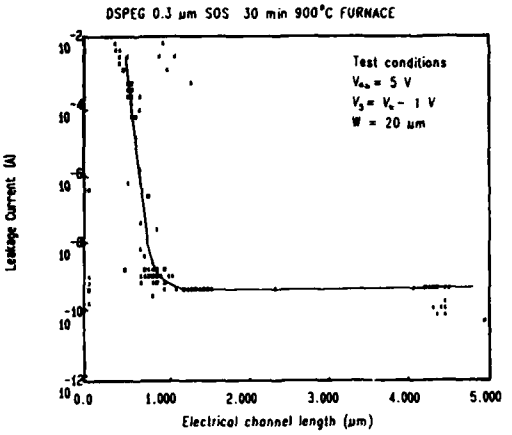


Figure 3

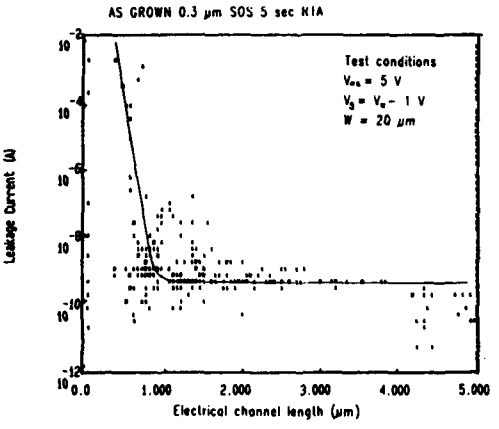


Figure 4

SILICON IC TECHNOLOGY USING COMPLEMENTARY MESFET'S

P A Tove, K E Bohlin, H Norde, U Magnusson, J Tirén, A Söderbärg,
M Rosling, F Masszi and J Nylander

Institute of Technology, Uppsala University,
Box 534, S-751 21 Uppsala, Sweden

We showed recently the successful operation of digital inverters and ring-oscillators using a new integrated circuit technology, complementary silicon MESFET's on silicon-on-sapphire, called CMES [1]. The technology involves fabricating normally-off p- and n-type MES-transistors by depositing two types of contact metals, one (Pt or Ir) having a high Schottky barrier ϕ_B on nSi and serving as gate on n-transistors and source and drain ohmic contacts on p-transistors, the other (Er or Tb) serving as high ϕ_B gate on pSi and ohmic S and D contacts for n-transistors. Figure 1 shows a cross-section of a CMES-inverter fabricated using SOS-technology, and its circuit diagram.

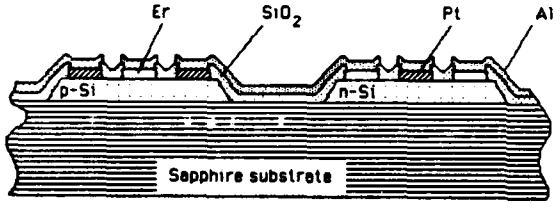


FIGURE 1a
Cross-section, inverter

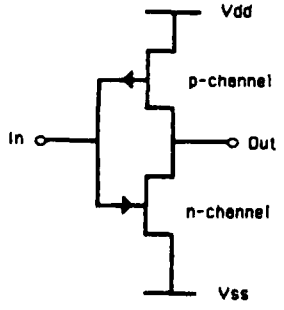


FIGURE 1b
Circuit-diagram, inverter

The main advantages of this technology compared to CMOS are 1) low power consumption (~factor 100 lower) because of low power supply voltage, typically <0.5 V necessitated by not driving the Schottky gates too much forward causing gate current; 2) high effective mobility and low noise because the carriers move deeper in the semiconductor than in the MOS case, thus avoiding interface and oxide states; 3) higher radiation resistance than for MOS (~factor 100) and 4) freedom from other MOS-oxide problems such as hot electron effects encountered at small dimensions.

Test chips containing n- and p-type transistors of different dimensions, complementary inverters and ring-oscillators, all using normally-off transistors have been fabricated. The threshold voltage V_T was 0.1 for both n- and p-type transistors (fig. 2) and figure 3 shows the transfer characteristics of a complementary inverter operating at V_{DS}=0.4 V. Measurements on a 7-stage ring-oscillator using transistors with 5 μm gate length show a gate-delay time of ~40 ns. According to the MESFET-theory [2], the gate-delay time is proportional to L^2. Computer simulations using a large-signal MESFET model [3] indicate a stage delay, t_D of 2 nsec for 2 μm gate length transistors with width ~50 μm, using V_T ~0.06 V and a supply voltage of 0.3 V.

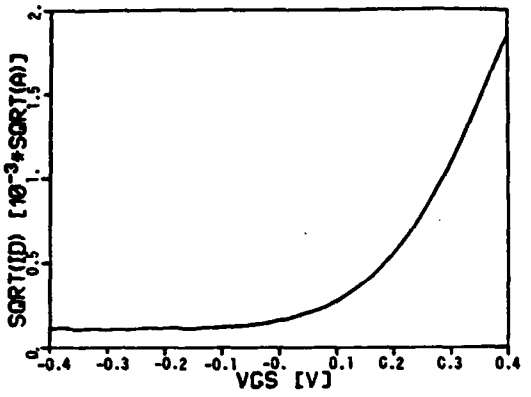


FIGURE 2
$\sqrt{I_D}$ versus V_G graph for n-channel transistor

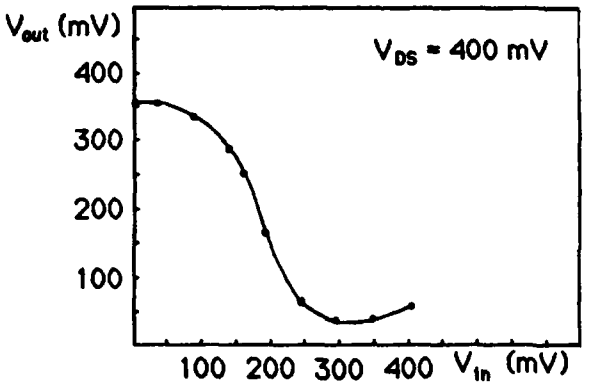


FIGURE 3
Transfer characteristic of inverter

Due to the complementary structure and similarities in processing, layout work is easily made by converting CMOS-layouts. New testchips are currently in process which contain more complex digital structures such as NAND, NOR, TG, XOR and static RAM cells. The layouts for these structures has been made using a modified CMOS CAD package. Analog circuits are also under development.

A special testchip containing a "fat FET" for characterization of the SOS materials has been produced [2]. This method is based on the study of drain current change, dI_D, occurring in the triode region when the channel width is incremented by Δw as the gate voltage is changed by $\Delta V_{GS}=V_{G1}-V_{G2}$, see figure 4.

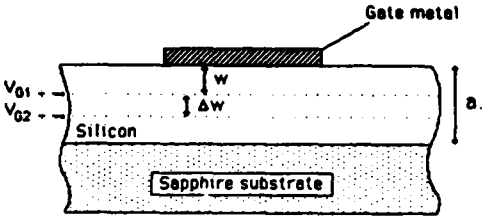


FIGURE 4
Principle of "fat-FET"

The mobility in the increment can be calculated using the relation

$$\Delta I_D = \mu E q N Z \Delta w \qquad (1)$$

where N is the doping concentration (=carrier concentration), E the field strength and Z the gate width. The channel increment is given by

$$\Delta w = \frac{\varepsilon_o \varepsilon_r \Delta V_G}{qNw} \qquad (2)$$

Now, we introduce the voltage drop due to I_D, in 2R, the resistances due to the space between gate and source (or drain) and neglect contact resistances (permissible if the barriers of the source and drain contacts are low enough). Using the relation

$$R = \frac{L}{q\mu NaZ}$$

where a is the film thickness and L the gate length, equal to the electrode spacing, we obtain

$$\mu = \frac{1}{V_{DS}}\left[\frac{L^2 g_m}{C_G} + \frac{2I_D L}{qNaZ}\right] \qquad (3)$$

where $C_G = \varepsilon_o \varepsilon_r \frac{ZL}{w}$ is the gate capacitance.

By measuring C_G as a function of V_G and coupling the obtained values with measured values of g_m for corresponding V_G values, the mobility of the carriers in the channel increment Δw can be determined.

The device used for these measurements were Si-MESFET transistors with L=0.05 mm and Z=2 mm which gives reasonable values of C_G.

Comparison with other methods of characterizing SOS, such as by optical methods [4] and by the spreading resistance technique [5] have also been made. Also, the change in transconductance, g_m, when MESFET's were exposed to ^{137}Cs and electron accelerator radiation (2 MeV) up to 10 Mrad has been studied. Our results thus far show a decrease in g_m of 15-35% for 3 measured transistors.

Other work describe silicon MESFET circuits withstanding radiation doses up to 100 Mrad [6]. This indicates a radiation hardness approximately 100 times higher than for MOS-devices why CMES-structures would be useful in high-radiation environment applications such as satellite, accelerator and nuclear safety systems. Also, the low power consumption suggests the use of CMES-technology for portable and other systems where very low power consumption is important.

REFERENCES

[1] Bohlin K, Tove P A, Magnusson U and Tirén J, El. Letters 23(5) (1987) 205.

[2] Tove P A, Bohlin K, Masszi F, Norde H, Nylander J, Tirén J, Magnusson U, Rosling M, Söderbärg A and Masszi N, UPTEC 86 131 R (1986).

[3] Nylander J O, Masszi F and Tove P A, to appear in COMPEL (Intern. Journ. for Computation etc); accepted May 1987.

[4] Grivitskas V, Willander M and Tellefsen J A, J Appl. Phys. 55(8) (1984) 3169.

[5] Lin A L, Maddox R L and Mee J E, J Appl. Phys. 57(6) (1985) 2091.

[6] Houston T W, Hite L R, Darley H M, Shedd W M, Zugich M H and Lapierre D C, Trans. Nucl. Sci. NS-31, 6 (1984) 1483.

Session C2.3

MOS Modelling II

Chairman: S. Selberherr

Tuesday, September 15, 1987

THE VOLTAGE-DOPING TRANSFORMATION A NEW APPROACH TO THE MODELLING OF MOSFET SHORT-CHANNEL EFFECTS

Tomasz SKOTNICKI*, Gérard MERCKEL and Thierry PEDRON

CNET-CNS, BP 98, Chemin du Vieux Chêne
38243 MEYLAN, FRANCE

In this paper we show that the influence of the drain-source field on the potential barrier height is physically equivalent and can be replaced by the reduction in channel doping concentration according to the following formula $N^*=N-2\varepsilon_s V^*_{DS}/qL^2$ derived from the 2-d Poisson equation. Thus the actual barrier height for any drain bias V_{DS} and channel length L can be easily calculated using the well-known 1-d (long-channel) solutions. This simple but general procedure, hereafter called the Voltage-Doping Transformation (VDT) has been examined with fairly good results by comparison of the analytically calculated potential distributions with 2-d numerical simulation. An examplary application of the VDT to threshold voltage calculations is also shown.

1. INTRODUCTION

The short-channel effects are one of the major constraints in VLSI MOSFETs miniaturization. In the case of analytical modelling of short-channel effects the main difficulty lies in a determining the potential barrier heights as they are strongly affected by two perpendicular gate and drain electrical fields. The 2-d character of short-channel effects leads to a tradeoff between the complexity of the model on the one hand and its validity and accuracy on the other hand.

The Voltage-Doping Transformation (VDT) proposed in this paper consists in replacing the influence of the lateral drain-source field by the equivalent reduction in channel doping concentration. It will be shown that the VDT enables the models of short-channel effects to be developed basing on assumptions of high degree of reality and without any need for numerical calculations.

2. VOLTAGE-DOPING TRANSFORMATION

In Fig. 1 the x-axis is the loci of the minimum of potential distribution in the y-direction and y is the distance from the source measured along a given current line.

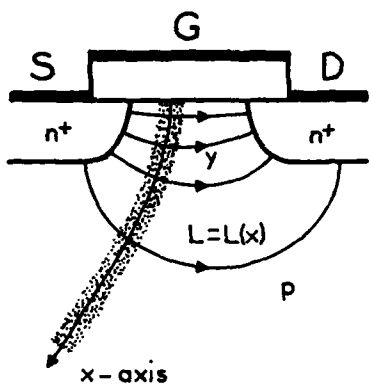


FIGURE 1

Domain of solution (dotted region) to the Poisson equation and a schematic illustration of current lines.

In the close vicinity of the x-axis (dotted region in Fig. 1) the Poisson equation can be rewritten in its conventional form :

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = \frac{q}{\varepsilon_s} N \qquad (1)$$

As discussed in [1] and [2] through numerical

* T. Skotnicki is also a member of the ITE-CEMI, Warsaw, Poland

simulation, a potential distribution along the channel in the short-channel case varies quadratically with the distance. Thus we assume that :

$$v(y)=ay^2+by+c \tag{2}$$

Lateral potential distribution v is not necessarily required to be equal to solution $\psi(x,y)$ to eqn. (1). Only an equality of second derivatives of these two functions is needed for our purposes. To approach this requirement we impose the following boundary conditions on function v :

- at the source end $v(0)=V_{bi}+V_{SB}$ (3a)
- at the drain end $v(L)=V_{bi}+V_{SB}+V_{DS}$ (3b)
- at the virtual cathode $\min[v(y)]=\phi_{vc}$ (3c)

where L is the current line length as shown in Fig. 1. In general $L(x)\geq L_{el}$ and $(L-L_{el})$ increases with x and decreases with x_j.

Coefficients a, b and c can be found directly from eqns (2) and (3), but here only a will be of interest because :

$$\frac{\partial^2\psi}{\partial y^2}=\frac{\partial^2 v}{\partial y^2}=2a=2\frac{V^*_{DS}}{L^2} \tag{4a}$$

where

$$V^*_{DS}=V_{DS}+2(V_{bi}+V_{SB}-\phi_{vc}) +2\sqrt{(V_{bi}+V_{SB}-\phi_{vc})(V_{DS}+V_{bi}+V_{SB}-\phi_{vc})} \tag{4b}$$

Substitution of eqn. (4) into eqn. (1) yields:

$$\frac{\partial^2\psi}{\partial x^2}=\frac{q}{\varepsilon_s}N^* \tag{5a}$$

where

$$N^*=N-\frac{2\varepsilon_s V^*_{DS}}{qL^2} \tag{5b}$$

The last equation is the essence of the Voltage-Doping Transformation which, as we have shown, enables the 2-d Poisson equation (1) to be reduced to the 1-d form given by eqn. (5a). Physically it means that the influence of the lateral drain-source field on the potential barrier height is equivalent to and can be replaced by the reduction in doping concentration according to eqn. (5b). A negative value of N^* indicates that all impurity ions are tied by a drain field which simply means that a MOSFET is in the punch-through mode.

3. THRESHOLD VOLTAGE MODEL

ϕ_{vc} is in fact the searched for potential distribution ($\phi_{vc}(x) \equiv \Psi(x,y=\text{virtual cathode})$ which appears to involve an iterative solution to eqn. (5a). Fortunately, for V_{th} calculations no iterations are necessary. Because of the curvature of the current lines, as shown in Fig. 1, L increases rapidly to infinity inside the bulk region which produces $N^* \simeq N$ in the bulk, irrespective of the V^*_{DS} value and consequently of the ϕ_{vc} value. On the other hand, in the surface region it is reasonable to assume $L(x)=L_{el}=\text{const}$ and $\phi_{vc}=\eta\phi_{svc}$ since, near the surface, $\phi_{vc}(x)$ values are close to ϕ_{svc}. In the last expression η is a spreading parameter accounting for the transverse distribution of $\phi_{vc}(x)$ values and $\phi_{svc}=\phi_{vc}(0)$. ϕ_{svc} is known and equal to Ψ_c where Ψ_c is the critical voltage needed for the onset of strong inversion (usually $\Psi_c=2\phi_{FB}+V_{SB}$). Thus we adopt here :

$$V^*_{DS}=V_{DS}+2(V_{bi}+V_{SB}-\eta\psi_c) +2\sqrt{(V_{bi}+V_{SB}-\eta\psi_c)(V_{DS}+V_{bi}+V_{SB}-\eta\psi_c)} \tag{6}$$

which removes the necessity of any iterative calculations.

With the use of V^*_{DS} as given by eqn. (6) we can solve eqn. (5a) for any arbitrary N(x) doping profile. However, for the sake of simplicity, we will focus our considerations on the box approximated profiles. The N_s and λ will denote the doping concentration and thickness of the near-surface box respectively, while N_B will

be the bulk doping concentration. We also assume that there is no bulk punch-through which means that $N^*_B \approx N_B$ and only the near-surface box concentration is reduced by the lateral drain-source field according to the relationship $N^*_s = N_s - 2\varepsilon_s V^*_{DS}/qL^2$, where V^*_{DS} is given by eqn. (6). As neither N^*_B nor N^*_s depends on x, the solution to eqn. (5a) poses no difficulty and results in the quadratic dependence of potential Ψ on distance x.

In Fig.2 the analytical solution for potential Ψ is compared with the results of the MINIMOS [3] simulation performed for a set of MOSFETs with different channel lengths.

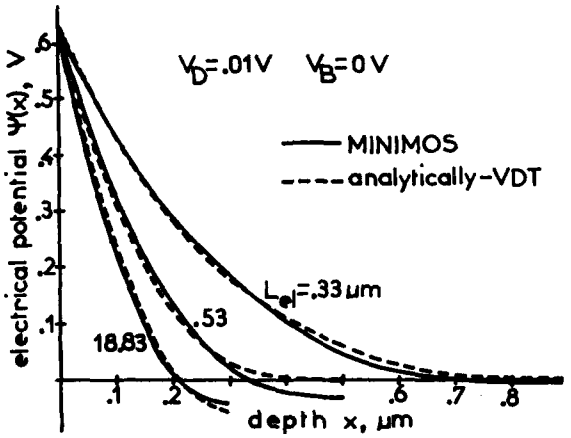


FIGURE 2

Plots of transverse potential distributions for implanted MOSFETs with different channel lengths. The parameters of the doping box approximation are : N_s=1E16cm^{-3}, λ=0.31µm, N_B=1E15cm^{-3}. The other parameters are as follows: T_{ox}=0.0700µm, x_j=0.72µm, V_{FB}=-1.023V and η=1.

It is worth noting that the VDT accurately accounts for the increase in depletion depth with channel shortening which, as seen in Fig.2, can double or even triple its long channel value, thus bringing into question the neglect of this effect as is the case for example in [4]. The close aggreement between the analytical and the numerical solutions, as seen in Fig. 2, indicates the correctness of the assumptions (2) and the VDT itself.

Having the solution for the electrical potential one can calculate the space charge Q_B tied by the gate field. Substitution of this Q_B in the well-known expresion for the long-channel threshold voltage yields :

- for $w_1 \leq \lambda^2$ and $N^*_s > 0$:

$$V_{th} = V_{FB} + \psi_c - V_{SB} + \frac{1}{C_{ox}} \sqrt{2q\varepsilon_s (\psi_c - \phi_0)\left(N_s - \frac{2\varepsilon_s V^*_{DS}}{qL_{el}^2}\right)} \qquad (7a)$$

- for $w_1 > \lambda^2$ or $N^*_s \leq 0$:

$$V_{th} = V_{FB} + \psi_c - V_{SB} + \frac{q}{C_{ox}}\lambda\left(N_s - N_B - \frac{2\varepsilon_s V^*_{DS}}{qL_{el}^2}\right) + \frac{q}{C_{ox}} \sqrt{\frac{2\varepsilon_s}{q} N_B \psi_c - N_B \lambda^2 \left(N_s - N_B - \frac{2\varepsilon_s V^*_{DS}}{qL_{el}^2}\right)} \qquad (7b)$$

where V^*_{DS} is given by eqn. (6), $w_1 = 2\varepsilon_s \psi / qN^*_s$ and $\phi_0 = -kT \ln(N^*_s/N_B)/q$.

The V_{th} model constituted by eqn. (7) shows an approximate linear dependence on V_{DS} bias, an approximate inverse-quadratic dependence on channel lenght L_{el} and an inverse dependence on oxide capacitance C_{ox} which is in good agreement with what has been found by Klaassen and de Groot in [5].

4. RESULTS

In Fig. 3 the threshold voltage variation with drain biases at different bulk biases equal to 0V (Fig. 3a) and -2V (Fig. 3b) is shown.

The calculated results closely agree with measured data taken on implanted MOSFETs with different channel lengths.

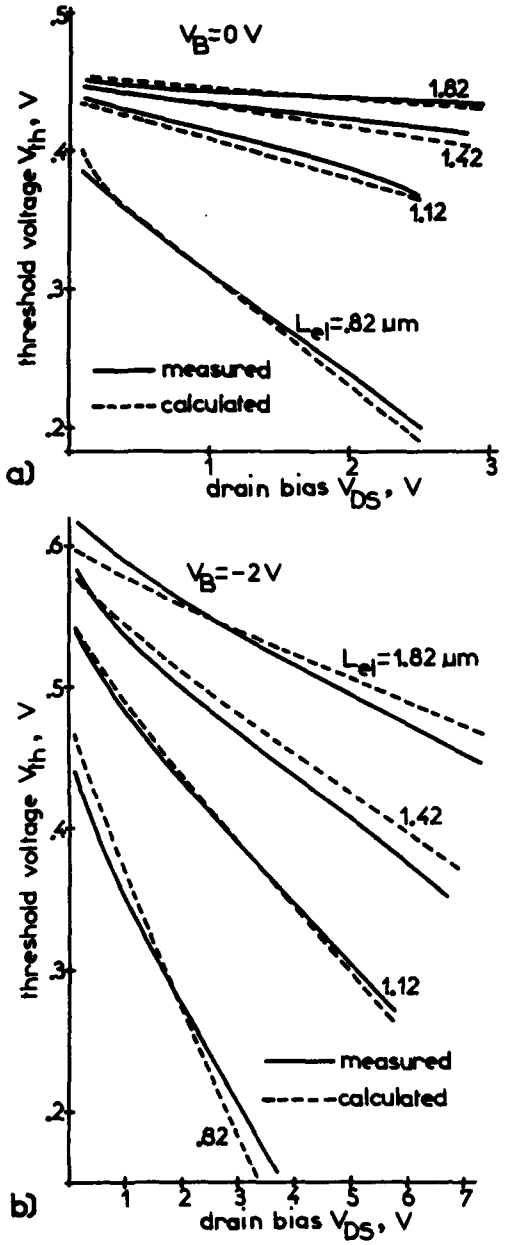


FIGURE 3

Threshold voltage versus drain bias at different bulk biases a) V_B=0 V and b) V_B=-2 V. Measurements have been taken on implanted MOSFETs with different channel lengths. The parameters of the doping box approximation are : N_a=1.6E16cm^{-3}, λ=0.29µm, N_B=4E14cm^{-3}. The other parameters read : T_{ox}=0.0250µm, x_j=0.27µm, V_{FB}= -0.706 V and η=0.942.

5. CONCLUSIONS

The Voltage-Doping Transformation has been derived and its validity has been shown. The new V_{th} model proposed here has aimed to show only an examplary application of the VDT. For this reason we have omitted such "secondary" effects as, for example, the dependence of the mean current line length on junction depth x_j thus neglecting the V_{th} dependence on x_j. Examination of the dependance of η on process and electrical parameters would also be required in order to avoid its fitting. We believe that this and other possible shortcomings will be improved in further applications of the VDT.

Nevertheless the new V_{th} model compares favourably with measured data taken on implanted short and long-channel MOSFETs. Consequently, the validity of the VDT as well as the new V_{th} model itself has been confirmed. The VDT is a very general tool and thereby is believed to be equally useful for V_{th} as well as punch-through current analytical modelling.

ACKNOWLEDGEMENT

The authors would like to express their thanks to Mr. C. Denat and Mr. A. Maitre from CNET-CNS for their help with measurements and data processing.

REFERENCES

[1] T. Skotnicki and W. Marciniak, Solid-St. Electron., vol.29, N°11, pp. 1115-1127, 1986.

[2] T. Skotnicki, Ph.D. thesis, Institute of Electron Technology CEMI, Warsaw, 1985.

[3] S. Selberherr, A. Schütz and H.W. Pötzl, IEEE Trans. Electron Devices, vol. ED-27, pp. 1540-1550, Aug. 1980.

[4] T. Toyabe and S. Asai, IEEE Trans. Electron Devices, vol. ED-26, pp. 453-461, Apr. 1979.

[5] F.M. Klaassen and W.C.J. de Groot, Solid-St. Electron., vol. 23, pp. 237-242, N°3, 1980.

A NOVEL REALISTIC MODEL FOR THRESHOLD VOLTAGE OF SHORT CHANNEL MOSFETS

M. Orlowski and Ch. Werner

SIEMENS Corporate Research and Development, Microelectronics
Otto-Hahn-Ring 6, 8000 München 83, FRG

The proposed model gives more realistic account of the physics of short channel behavior than its predecessors allowing fits to experimental data and simulated results without resorting to unrealistic values for clearly defined parameters. The most distinct and new feature of this model is the prediction of the onset of the threshold voltage reduction at a critical range of channel lengths. The model is formulated in terms of simple analytic formulae.

INTRODUCTION

We present a novel realistic model for threshold voltage, V_{Th}, which for the first time gives a proper physical description of the mechanism leading to threshold voltage reduction for short channels. In particular and in contrast to extant models /1,2,3/ this model predicts a critical value of channel length below which the short channel behavior comes into effect. Moreover this model posseses two fit parameters with clear physical meaning which allow to fit V_{Th} as a function of the channel length L, even for curves obtained by 2D device simulation where the subdiffusion length, y_{sub}, junction depth x_j, and the effective channel doping n_{eff}, are unambiguously defined. As all of its predecessors this model is rooted in Yau's charge sharing model /1/. However, despite the simplicity of the charge sharing idea the models hitherto ran into conceptual inconsistencies when attempting to incorporate the charge sharing idea into a sensible model to predict threshold voltage. Clearly - adhering to the principle of the weakest point in the hose - for sufficiently long channels V_{Th} must stay fixed rather than to approach "long channel V_{Th}" asymptotically /2/.

MODEL

The basic idea of the present model is illustrated in Fig.1. Consider first the depletion zones generated by the gate electrode, source or drain region separately. Superposing the three pieces in purely geometrical way to form a MOSFET structure, we obtain overlaps of the depletion zones at the source and drain sides (Fig.1a). This geometric construction, however, is in conflict with the Poisson's equation, since the charge in the overlap region is insufficient to terminate the field lines originating from source/drain and gate charges. Somehow additional charges have to be provided to restore the overall charge neutrality. Assuming for convenience a uniform channel doping, exactly the same area of the overlap has to be supplemented beyond the boundaries of the depletion zone shown in Fig.1a. As a first order geometric approximation this area is supposed to be distributed over the area of a triangle as shown in Fig.1b. If the channel doping is nonuniform, the two pertinent areas are not equal, in general, and a scaling factor $\alpha_{s/d}$ has to be introduced to account for the channel nonuniformity. Of course, more sophisticated geometries may be employed to approach more accurately the actual shape of the effective depletion zone.

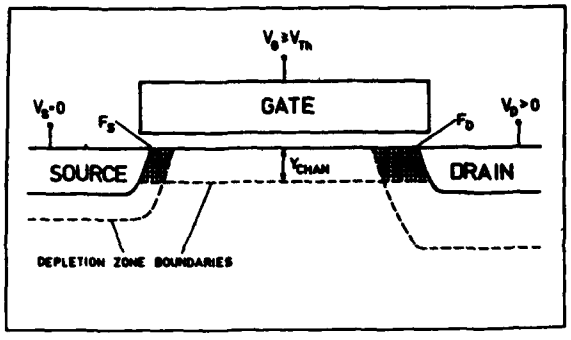


FIGURE 1a

Geometric overlap of the depletion generated by the gate, source and drain independently from one another.

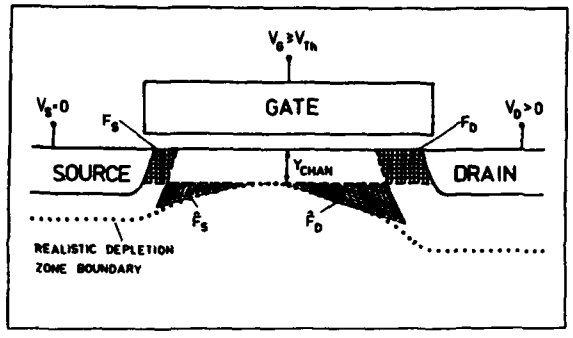


FIGURE 1b

For the shaded areas it holds: $\hat{F}_S = \alpha_S F_S$ and $\hat{F}_D = \alpha_D F_D$. For uniform channel doping $\alpha_S = 1 = \alpha_D$.

A simple construction of the geometry $\hat{F}_D$ (a triangle) is demonstrated in Fig.2. The given quantities are x_j, y_{chan}, and y_D, the latter two depending on drain V_D and substrate V_{sub} bias in the well-known manner. One calculates then

$$x_D = ((x_j^2 + 2x_j y_D)^{1/2} - x_j)*\beta ,$$

$$x_{oD} = x_j + y_D - x_D = r_D - x_D, \quad B=(x_{oD}, y_{chan}),$$

$$S = ((r_D^2 + y_{chan}^2)^{1/2}, y_{chan}),$$

$$P = (r_D/(1+(\gamma y_{chan}/x_{oD})^2)^{1/2},$$
$$\gamma y_{chan}/x_{oD} * r_D/(1+(\gamma y_{chan}/x_{oD})^2)^{1/2}),$$

and $$F_D = 1/2*(y_{chan}(r_D^2+y_{chan}^2)^{1/2} + r_D^2 \arcsin(y_{chan}/r_D)) - x_{oD}*y_{chan} .$$

Finally the lateral coordinate t of the point $T=(t,y_{chan})$ is determined such that the area $\hat{F}_D$ of the triangle PST is equal to $F_D\alpha_D$. The above construction applies to both, source and drain side. Now: as long as the triangles from both channel ends do not overlap, V_{Th} remains unaltered. If they do overlap for sufficiently short channels the threshold voltage will be reduced. In determining its reduction we follow the usual procedure, calculating the charges, Q_{corr}, in regions where the triangle overlap over the range $L_{over} \leq 2L_D$, L_D being the Debey length.

Next, we subtract this charge from the space charge supplied by the gate electrode according to eq.1 :

$$Q_{eff} = q*n_{eff}*y_{chan}*W*L_{over} - Q_{corr} \qquad (1)$$

The threshold voltage is then given by

$$V_{Th} = U_{FB} + 2kT/q*\ln(n_{eff}/n_i) + d_{ox}Q_{eff}/(\varepsilon_{ox}WL_{over}). \qquad (2)$$

A more detailed account of this model and its refined versions will be given elsewhere /4/.

In this model the reduction of the lateral thickness of the space charge region due to the curvature of the source/drain pn junctions are considered by the formula given above for x_D. To achieve a better adjustment to the actual overlap of the space charge regions a fit parameter β has been introduced . A second fit parameter γ determines the flatness of the triangles approximating the actual geometry of

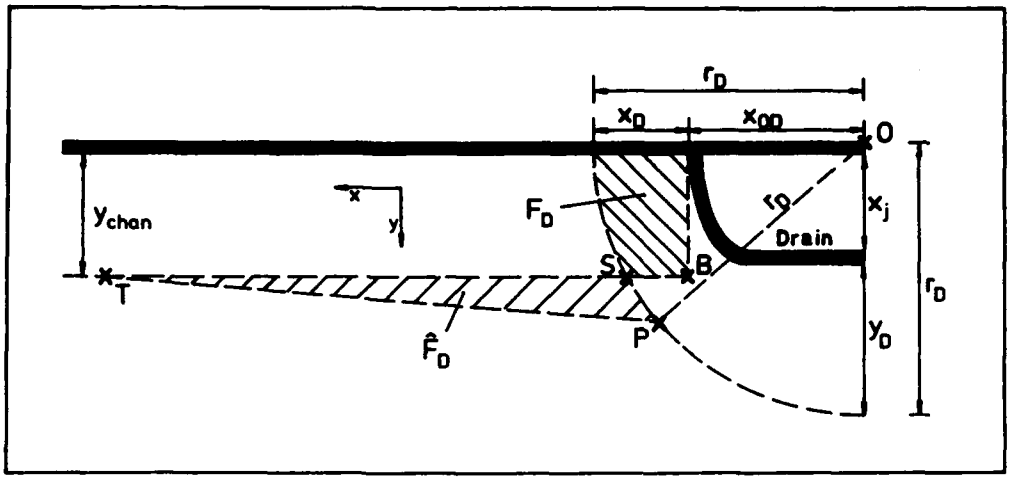


FIGURE 2

Schematic diagram for the construction of the triangle PST.

the depletion zone. The flatter the triangle the larger its lateral extent. Therefore this parameter determines the critical channel length at which the short channel effects start to be effective. It was found that a proper adjustment of both parameters and the nonuniformity scaling factor allow to fit experimental or simulated curves using realistic or actual values for y_{sub}, x_j, and n_{eff}. The formulae of the model are lengthy but sufficiently simple to be used in such programs as SPICE.

APPLICATIONS

The model describes correctly the dependence of the threshold voltage as a function of the channel length for a variety of parameters such as V_D, V_{sub}, x_j, n_{eff}, and oxide thickness d_{ox}. In Fig.3 as an example typical curves are shown for different drain and substrate biases.

We like also to mention that this model based on the concept expounded in Fig.1 and 2 can be hybridized with specific approaches discussed in the vast literature on threshold voltage models including the narrow width effects.

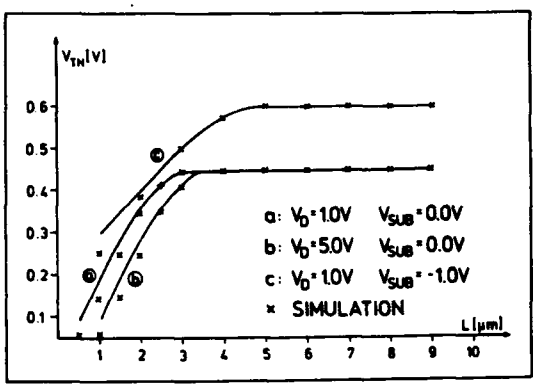


FIGURE 3

Typical curves of the threshold voltage as a function of the channel length. Note that for sufficiently large L the curves remain flat.

ACKNOWLEDGEMENTS

We would like to thank Dr. Jäntsch for helpfull discussions.

REFERENCES

/1/ Yau, L. D., Solid-St. Electron. 17 (1974) 1059-1063

/2/ Akers, L. A. and Sanchez,J.J.,Solid -St. Electron. 25 (1982) 621-641

/3/ Jäntsch, O.,Solid-St. Electron. 25 (1982) 59-61

/4/ Orlowski, M. and Werner, Ch. to be published

SENSITIVITY ANALYSIS FOR DEVICE DESIGN

A. Gnudi, P. Ciampolini, R. Guerrieri, M. Rudan and G. Baccarani

Dipartimento di Elettronica, Università di Bologna
viale Risorgimento 2, 40136 Bologna, Italy

Abstract

In this paper we propose a sensitivity-analysis technique for device design. By this method, we determine the linearized variations of the device terminal characteristics following some change either in the impurity distribution, or in device geometry, such as channel length and oxide thickness. This technique has been implemented in our general-purpose two-dimensional device-analysis program (HFIELDS) and proved to be very efficient, as only the assembly of the RHS and one back-substitution is required in order to achieve the final result. It is believed that the present method can be profitably used for both deterministic and statistical device design.

1. Introduction

As device miniaturization progresses towards submicron feature sizes, fabrication tolerances cannot be made to scale in direct proportion with the device physical dimensions. As a consequence, circuit designers must live with an increased spread in device terminal characteristics which, if not properly accounted for, can lead to design marginality and to a significant degradation of the fabrication yield [1]. Two typical examples are *i)* the lightly-doped drain extensions (LDD), which help reducing the lateral electric field at the drain end of the channel but, at the same time, inevitably introduce an increased parasitic series resistance which adversely affects the device transconductance, and *ii)* the increase in the substrate impurity concentration, which improves the device performance from the standpoint of short-channel effects, but leads to an increased parasitic junction capacitance and to a more severe threshold sensitivity to substrate potential.

Due to the insufficient accuracy of analytical models as predictive tools of device performance, numerical device-analysis programs are currently being used for device design and optimization. These programs, however, require large CPU resources in view of the inherent nonlinearity of the fundamental device equations, and to the large number of grid point necessary for an adequate device description in two and three dimensions. Hence, repeated simulation of the device under investigation for a variety of geometry and impurity profiles is an expensive and time-consuming task.

In order to alleviate the above problem, we have developed a steady-state sensitivity analysis technique which provides the linearized variations of the device terminal characteristics following some change in device geometry and/or impurity profiles. As sensitivity computation relies on linearization of the fundamental semiconductor equations, the overhead of a sensitivity analysis superimposed on the simulation of the nominal device is negligible and, in our belief, can be of great help to the device designer both for design centering purposes and in order to assess the possible spread of the resulting electrical characteristics.

2. Sensitivity Analysis

The starting point for the sensitivity analysis is the solution of the fundamental semiconductor equations in steady state

$$\operatorname{div}\mathbf{D} = q(p - n + N_D - N_A) \qquad (2.1,\text{a})$$

$$\operatorname{div}\mathbf{J}_n = q(R - G) \qquad (2.1,\text{b})$$

$$\operatorname{div}\mathbf{J}_p = -q(R - G) \qquad (2.1,\text{c})$$

representing Poisson equation (2.1,a) and carrier continuity equations for electrons (2.1,b) and holes (2.1,c). The latter are supplemented by the constitutive equations

$$\mathbf{D} = -\epsilon_s \operatorname{grad}\varphi \qquad (2.2,\text{a})$$

$$\mathbf{J}_n = -q\mu_n n \operatorname{grad}\varphi + qD_n \operatorname{grad} n \qquad (2.2,\text{b})$$

$$\mathbf{J}_p = -q\mu_p p \operatorname{grad}\varphi - qD_p \operatorname{grad} p \qquad (2.2,\text{c})$$

where the symbols in eqs. (2.1), (2.2) have the customary meaning of semiconductor theory. The solution of the equations in two dimensions is carried out in program HFIELDS (*Hybrid FInite ELement Device Simulator*), on a triangular grid, by means of the so-called *Box Integration Method* [2]. The discretization of the continuity equations is carried out by means of a two-dimensional generalization of the Gummel-Scharfetter

method. The assembly of the coefficient matrix is done on a triangle-by-triangle basis, which allows for a general and simple management of interfaces and boundary conditions. The procedure transforms eqs. (2.1) into the following set of non-linear algebraic systems

$$\epsilon_s \sum_{j \neq i} \frac{d_{ij}}{s_{ij}} (\varphi_i - \varphi_j) + q \Omega_i (n_i - p_i - N_i) = 0 \quad (2.3,\text{a})$$

$$D_n \sum_{j \neq i} \left[\frac{d_{ij}}{s_{ij}} (B_{ji} n_i - B_{ij} n_j) \right] + \Omega_i U_i = 0 \quad (2.3,\text{b})$$

$$D_p \sum_{j \neq i} \left[\frac{d_{ij}}{s_{ij}} (B_{ij} p_i - B_{ji} p_j) \right] + \Omega_i U_i = 0 \quad (2.3,\text{c})$$

where $B_{ij} = (u_j - u_i)/[\exp(u_j - u_i) - 1]$ is the Bernoulli function ($u = q\varphi/kT$ being the normalized potential), and suffix i indicates the i^{th} node of the grid. The sums are extended over all the nearest neighbours j of node i.

It is important to notice that, in eqs. (2.3), the geometrical factors are the box area Ω_i and the ratio $A_{ij} = d_{ij}/s_{ij}$, where s_{ij} is the length of the element side S_{ij} connecting nodes i and j, and d_{ij} is the cross-section, i.e. the distance of the circumcenter of the triangular element from S_{ij}. Eqs. (2.3) also contain the nodal values of the donor and acceptor impurity concentrations N_D, N_A, which appear explicitly in eq. (2.3,a), and implicitly in eq. (2.3,b), (2.3,c) via the carrier mobility. Thus, after dropping the indices, we may rewrite eqs. (2.3) as

$$F_\varphi(\varphi, n, p; N_D, N_A, \Omega, A) = 0 \quad (2.4,\text{a})$$

$$F_n(\varphi, n, p; N_D, N_A, \Omega, A) = 0 \quad (2.4,\text{b})$$

$$F_p(\varphi, n, p; N_D, N_A, \Omega, A) = 0 \quad (2.4,\text{c})$$

If the grid contains N_c nodes and N_s element sides (excluding those pertaining to ohmic contacts) φ, n, p, N_D, N_A and Ω are N_c-dimensional vectors, while A is an N_s-dimensional vector.

We now assume that, due to some change in the process parameters, the nominal values of the donor and acceptor densities $N_D^{(o)}$ and $N_A^{(o)}$ are modified by a small amount, i.e.

$$N_D = N_D^{(o)} + \delta N_D \quad (2.5,\text{a})$$

$$N_A = N_A^{(o)} + \delta N_A \quad (2.5,\text{b})$$

As a consequence, the electric potential and carrier concentrations will experience a small change

$$\varphi = \varphi^{(o)} + \delta\varphi \quad (2.6,\text{a})$$

$$n = n^{(o)} + \delta n \quad (2.6,\text{b})$$

$$p = p^{(o)} + \delta p \quad (2.6,\text{c})$$

such that eqs. (2.4) are again fulfilled. Differentiating these equations leads to the system

$$\begin{bmatrix} \left(\frac{\partial F_\varphi}{\partial \varphi}\right)_0 & \left(\frac{\partial F_\varphi}{\partial n}\right)_0 & \left(\frac{\partial F_\varphi}{\partial p}\right)_0 \\ \left(\frac{\partial F_n}{\partial \varphi}\right)_0 & \left(\frac{\partial F_n}{\partial n}\right)_0 & \left(\frac{\partial F_n}{\partial p}\right)_0 \\ \left(\frac{\partial F_p}{\partial \varphi}\right)_0 & \left(\frac{\partial F_p}{\partial n}\right)_0 & \left(\frac{\partial F_p}{\partial p}\right)_0 \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta n \\ \delta p \end{bmatrix} =$$

$$= - \begin{bmatrix} \left(\frac{\partial F_\varphi}{\partial N_D}\right)_0 \\ \left(\frac{\partial F_n}{\partial N_D}\right)_0 \\ \left(\frac{\partial F_p}{\partial N_D}\right)_0 \end{bmatrix} \delta N_D - \begin{bmatrix} \left(\frac{\partial F_\varphi}{\partial N_A}\right)_0 \\ \left(\frac{\partial F_n}{\partial N_A}\right)_0 \\ \left(\frac{\partial F_p}{\partial N_A}\right)_0 \end{bmatrix} \delta N_A \quad (2.7)$$

where the coefficient matrix at the LHS is the jacobian matrix of the F functions. It should be noticed that, in order to solve eqs. (2.4) in the nominal conditions, a Newton-Raphson method is used, this requiring the assembly and factorization of the jacobian matrix. Thus, the overhead associated with the sensitivity analysis is simply given by the assembly of the RHS and a back substitution, requiring very little additional effort.

We now assume that, due to process parameter changes, the geometry of the device under investigation is somewhat modified. Therefore, the grid points will have to move in order to accomodate the above geometrical changes and to conform to the new device geometry. Thus we have

$$x_i = x_i^{(o)} + \delta x_i \quad (2.8,\text{a})$$

$$y_i = y_i^{(o)} + \delta y_i \,. \quad (2.8,\text{b})$$

It is quite obvious that the choice of δx_i and δy_i is not unique. However, we stress that such a choice is not critical as far as the integral physical quantities, such as terminal currents, are concerned. From a practical standpoint, we draw two vertical or horizontal straight lines crossing the device cross-section, and we stretch or compress uniformly by a predefined amount the portion of the device lying between those two lines. The grid points are therefore moved horizontally or vertically in direct proportion to their distance from one of the two lines. Of course, one of the portions of the device lying outside the stretched or compressed band will have to be shifted uniformly.

The change in coordinate values will modify the geometrical parameters Ω_i and A_{ij} as follows

$$\Omega_i = \Omega_i^{(o)} + \sum_{k=1}^{N_c} \left(\frac{\partial \Omega_i}{\partial x_k} \delta x_k + \frac{\partial \Omega_i}{\partial y_k} \delta y_k \right) \quad (2.9,\text{a})$$

$$A_{ij} = A_{ij}^{(o)} + \sum_{k=1}^{N_c} \left(\frac{\partial A_{ij}}{\partial x_k} \delta x_k + \frac{\partial A_{ij}}{\partial y_k} \delta y_k \right) \,. \quad (2.9,\text{b})$$

The electric potential and the carrier concentration will then undergo some changes $\delta\varphi$, δn and δp such that eqs. (2.4) are again fulfilled. Thus we find the following set of linear equations

$$\begin{bmatrix} \left(\frac{\partial F_\varphi}{\partial \varphi}\right)_0 & \left(\frac{\partial F_\varphi}{\partial n}\right)_0 & \left(\frac{\partial F_\varphi}{\partial p}\right)_0 \\ \left(\frac{\partial F_n}{\partial \varphi}\right)_0 & \left(\frac{\partial F_n}{\partial n}\right)_0 & \left(\frac{\partial F_n}{\partial p}\right)_0 \\ \left(\frac{\partial F_p}{\partial \varphi}\right)_0 & \left(\frac{\partial F_p}{\partial n}\right)_0 & \left(\frac{\partial F_p}{\partial p}\right)_0 \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta n \\ \delta p \end{bmatrix} =$$

$$= - \begin{bmatrix} \left(\frac{\partial F_\varphi}{\partial \Omega}\right)_0 \\ \left(\frac{\partial F_n}{\partial \Omega}\right)_0 \\ \left(\frac{\partial F_p}{\partial \Omega}\right)_0 \end{bmatrix} \delta\Omega - \begin{bmatrix} \left(\frac{\partial F_\varphi}{\partial A}\right)_0 \\ \left(\frac{\partial F_n}{\partial A}\right)_0 \\ \left(\frac{\partial F_p}{\partial A}\right)_0 \end{bmatrix} \delta A \quad (2.10)$$

and once again the coefficient matrix on the LHS is the jacobian matrix of the F functions, which is already available and factorized. Therefore, as for the previous case, sensitivity analysis following changes in device geometry only requires the assembly of the RHS and one back substitution.

Having determined the incremental quantities $\delta\varphi$, δn and δp, the terminal currents are found by a straightforward differentiation of the current densities at the contacts.

3. Numerical Results

In order to test the sensitivity analysis performed by HFIELDS, a MOSFET has been simulated. The latter was a rather typical 1.0 μm device, having an oxide thickness of 20 nm. The channel implant has been designed so as to set the threshold voltage at 0.7 V, resulting in a peak impurity concentration $N_o = 1.6 \times 10^{17}$ cm^{-3} and in a standard deviation $s = 0.076\,\mu$m. The nominal turn-on characteristic is illustrated in fig. 1 by the thick line labelled $N_o = 1.6 \times 10^{17}$. The peak impurity concentration was then raised by 30% to the value of 2.08×10^{17} cm^{-3}, leading to the linear approximation represented by the thin line, and to the exact non-linear solution represented by the thick line labelled 2.08×10^{17}. The exact and the linearized solutions exibit a reasonable agreeement above threshold, where they are nearly parallel and slightly displaced. One can observe that, as a consequence of the increased implantation dose, the threshold voltage is substantially increased, and the gain factor is appreciably reduced. The latter effect is due to mobility degradation, which occurs as a consequence of the increased average inversion-layer field for a given gate voltage. The Yamaguchi model [3] employed in this simulation fully accounts for the above effect.

The linearized solution exibits a dip in the subthreshold region due to the exponential dependence of drain current vs threshold voltage (and therefore implantation dose). As the voltage shift is several tenths of a volt and the current derivative is negative and large, the incremented current turns out to be negative. Thus, some care must be used when interpreting linearised results in rapidly-varying regions, unless appropriately-small increments are considered.

A similar procedure was followed for investigating the effect of the channel length. Fig. 2 shows the nominal turn-on characteristic, labelled $L = 1.0\,\mu$m, and the corresponding curves obtained with a 30% increase of the channel length. Once again, the exact and the linearized result, while slightly differing from a quantitative point of view, exhibit the same general behavior. The most important effect of a channel-length variation is a change in the device gain factor and transconductance, and a small increase in threshold voltage related to the short-channel effect.

Finally, we have investigated the effect of the oxide thickness on the MOSFET turn-on characteristics. As shown in fig. 3, the nominal oxide thickness of 20 nm has been raised by 30%, leading to the linear approximation (thin solid line) and to the exact non-linear solution (thick solid line) labelled $t_{ox} = 0.026\,\mu$m. As is well known, the oxide thickness affects at the same time the threshold voltage and the gain factor: hence the turn-on characteristic is shifted towards higher gate voltages, and exhibits a reduced slope in strong inversion. The discrepancy between the linear approximation and the exact result is a consequence of the nonlinarity of the current expression against oxide thickness. A final remark regards the dip of the linearized solution: once again this can be traced back to the exponential dependence of the subthreshold current upon threshold voltage which, in turn, is linearly affected by the oxide thickness. As already mentioned, the validity range of a linearized solution is considerably smaller when the quantity of interest is a strong function of the parameter under investigation.

4. Conclusions

In this paper we have discussed a sensitivity analysis technique implemented in our general-purpose device simulator. By performing a straightforward differentiation of the fundamental semiconductor equations, we can determine the sensitivity of the device to either geometrical changes, such as channel length and oxide thickness in MOSFET's, or to impurity-profile changes.

In order to exploit the full advantages of this technique, an integrated process and device analysis system ought to be used. The former should provide, in addition to the nominal geometry and impurity profiles, the predicted changes of the device physical structure following a variation in some process parameter. The latter should instead use such information to predict the final device sensitivity to the investigated process parameter. It should be mentioned that, being a pro-

cess simulator intrinsically time dependent, a transient sensitivity analysis would be required there. Such a transient sensitivity computation has already been implemented in a circuit simulator [4] using the direct differentiation approach, and could, in principle, be extended to process simulators.

The main advantages of such a technique are that one can combine the accuracy of numerical process and device simulators (as opposed to analytical ones) with the computational efficiency required when a large number of process parameters are to be investigated. The additional burden associated with sensitivity analysis on top of device simulation is only given by the assembly of the RHS and one back substitution of the factorized matrix. Thus the proposed technique is highly efficient and allows for a large number of process parameters to be investigated for each data point of the nominal dc characteristic.

Acknowledgements

This work has been supported by the Consiglio Nazionale delle Ricerche under the Project: "Materiali e Componenti per l'Elettronica a Stato Solido". Support from SGS-Microelettronica is also gratefully acknowledged.

REFERENCES

[1] S. R. Nassif, A. J. Strojwas, and S. W. Director: "FABRICS II: A Statistically Based IC Fabrication Process Simulator", *IEEE Trans. on Computer-Aided Design of ICAS*, vol. CAD-3, pp. 40-46, Jan. 1984.

[2] G. Baccarani, R. Guerrieri, P. Ciampolini and M. Rudan: "HFIELDS: a Highly-Flexible 2-D Semiconductor-Device Analysis Program", *Proceedings of the Fourth International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuits*, J. J. H. Miller Ed., Dublin: Boole Press, pp. 3–12, 1985.

[3] K. Yamaguchi, "A Mobility Model for Carriers in the MOS Inversion Layer", *IEEE Trans. on Electron Devices*, vol. ED-30, pp. 658–663, 1983.

[4] D. E. Hocevar, P. Yang, T. N. Trick and B. D. Epler: "Transient Sensitivity Computation for MOSFET Circuits", *IEEE Trans. on Computer-Aided Design of ICAS*, vol. CAD-4, pp. 609–620, Oct. 1985.

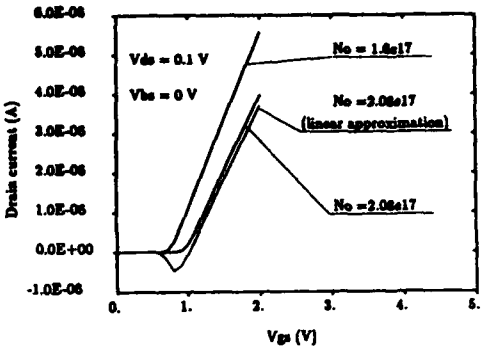


Fig. 1: Plot of the MOSFET turn-on characteristic in the nominal case where the peak impurity concentration of the channel implant was $N_o = 1.6 \times 10^{17}$ cm^{-3}. The thin and thick lines labelled $N_o = 2.08 \times 10^{17}$ cm^{-3} represent the linearised and the exact solutions, respectively, corresponding to an increase of the above parameter by 30%.

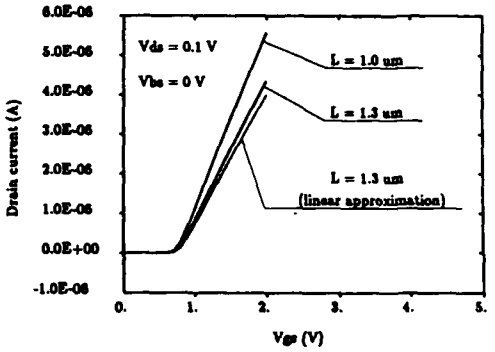


Fig. 2: Plot of the MOSFET turn-on characteristic in the nominal case where the net channel length was $L = 1.0\,\mu m$. The thin and thick lines labelled $L = 1.3\,\mu m$ represent the linearised and the exact solutions, respectively, corresponding to an increase of the above parameter by 30%.

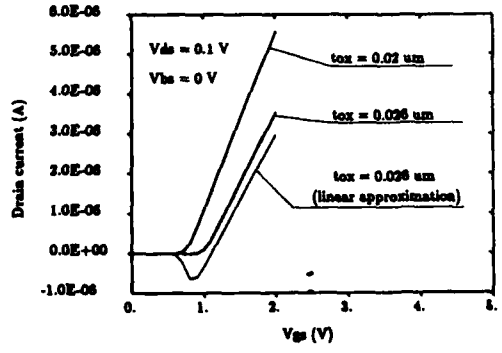


Fig. 3: Plot of the MOSFET turn-on characteristic in the nominal case where the oxide thickness was $t_{ox} = 20$ nm. The thin and thick lines labelled $t_{ox} = 26$ nm represent the linearised and the exact solutions, respectively, corresponding to an increase of the above parameter by 30%.

Session D2.2

Gettering II

Chairman: W. Orr-Arienzo

Tuesday, September 15, 1987

Gettering of Metal Precipitates

Gary Bronner[1] and James Plummer[2]

[1]IBM TJ Watson Research Center, PO Box 218, Yorktown Heights, NY 10598
[2]Center for Integrated Systems, Stanford University, Stanford, Ca. 94305

In this paper we deal with the effect of silicon interstitials on metal precipitates. We show that excess silicon interstitials which are injected by common gettering treatments cause most metal precipitates to shrink but a few to grow. Data from the literature confirms that the precipitates which grow by the absorption of silicon interstitials (primarily $FeSi_2$ and $NiSi_2$) are found exclusively in regions that act as net injectors of silicon interstitials. Other precipitates shrink in the presence of excess silicon interstitials and are easily gettered. This explains why silicon interstitial injection is essential for effective gettering.

Introduction

In earlier papers[1, 2] we analyzed the gettering of gold in silicon and reached the conclusion that several common gettering processes, including mechanical damage, argon implantation, high concentration phosphorus diffusion, and oxygen precipitation served as sources of silicon interstitials. Schematically, the role of silicon interstitials in the gettering of gold is illustrated in Figure 1. The important facts for gold are first that most gold is substitutional in silicon as opposed to being interstitial, i.e. $C^{eq}_{Au_{sub}} \gg C^{eq}_{Au_{int}}$ Additionally, the diffusion of gold is much faster as an interstitial atom, i.e. $D^{Au}_{int} \gg D^{Au}_{sub}$. For gold to diffuse it must be driven from its normal substitional site which occurs via the kick-out reaction

$$Si_{int} + Au_{sub} \Longleftrightarrow Au_{int}$$

Thus we see in Figure 1, silicon interstitials diffuse from the backsurface getter into the wafer. Near the backsurface, where the concentration of silicon interstitials is large, gold has been driven interstitial and then rapidly gettered. Away from the backsurface the gold remains immobile. As the silicon interstitials diffuse through the wafer, the profile of gettering gold exactly follows their movement. This behavior is expected and seen for all metals that have substantial substitutional concentrations, namely gold and platinum[3, 4].

Other metals which have little or no substitutional components such as copper, iron, and nickel would not depend on this mechanism for effective gettering. Based on the diffusion coefficients for interstitial metals in Figure 2, one would expect that metals could diffuse through a wafer in minutes even at temperatures as low as 600°C. Yet experimentally, it is

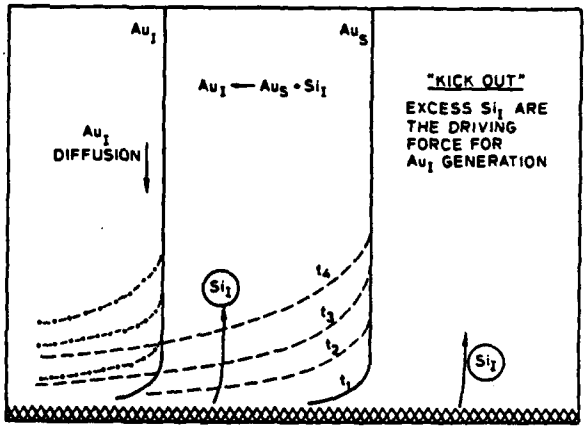


Figure 1. Schematic Representation of the Gettering of Gold.

found that gettering treatments that do not inject silicon interstitials, such as polysilicon are ineffective getters at low temperatures[5], which is surprising since the most common contaminants found in a processed silicon wafer are iron, copper, and nickel[6]. In this paper we propose that silicon interstitials are important for gettering the precipitates of these common metals. We will review the literature to present evidence that the most common device degradation is due to metal precipitates and then develop a theory to explain the role of silicon interstitials in the gettering of metal precipitates. It will be seen that the theory explains which metal precipitates can be gettered and which can not and that it agrees with microscopic evidence of metal precipitate crystalline structure.

Effect of Metal Precipitates on Devices

As early as 1960, Goetzberger and Shockley[7] recognized that metal precipitates play an important role in determining the breakdown and leakage characteristics of junction diodes. Devices intentionally contaminated with copper, iron, or manganese exhibit "soft" breakdown characteristics. Typically, the reverse leakage current has the form

$$I_{rev} \propto V^n \quad n=3 \text{ to } 7$$

They attributed this behavior to the presence of metal precipitates.

More recently, Busta and Waggener[8] using diodes and Lin et. al.[9] looking at breakdown in thin oxides confirmed these findings. For samples contaminated with copper, iron, nickel, tin, or zinc the degradation in device performance was due almost entirely to metal precipitates. Diodes exhibited soft leakage currents and capacitors showed breakdown a local hot spots associated with metal precipitates.

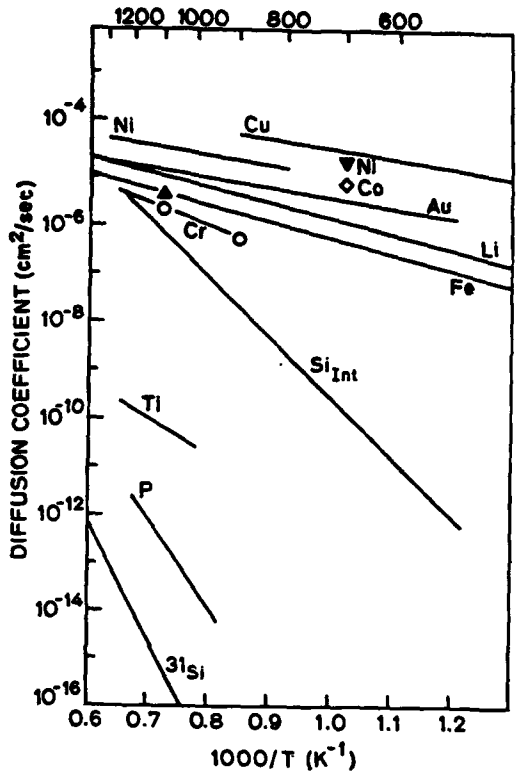


Figure 2. Diffusion Coefficients of Interstitial Metals.

Effect of Silicon Interstitials on Metal Precipitates

There has been considerable study of the microstructure of metal precipitates in silicon, most notably the work of Das[10] on copper and Vanderwalker[11] on nickel. The important conclusion of these studies was that when these metals precipitate they form metal silicides with *CuSi* being the only phase observed for copper, while nickel precipitated out as Ni_2Si, *NiSi*, and Ni_3Si depending on the temperature of anneal. The question of how the injection of silicon interstitials will affect these precipitates can be answered by considering the volumetric density of the precipitates compared to the silicon lattice. If a metal precipitate displaces more silicon atoms than it consumes (is a net injector of silicon interstitials), intro-

ducing silicon interstitials from a gettering region will tend to inhibit precipitation and dissolve existing precipitates. In more rigorous form, the interaction between the silicon lattice, interstitial metal, precipitated metal, and silicon interstitials can be written

$$Si_{lattice} + M_{int} \Longleftrightarrow M_xSi_{y_{prec}} + \Gamma Si_{int}$$

with

$$\Gamma = \frac{y}{x}\left(\frac{T_{Si}}{T_{M_xSi_y}} - 1\right)$$

where T_{Si} is the density of silicon atoms in the silicon lattice and $T_{M_xSi_y}$ is the density of silicon atoms in the metal precipitate. By balancing the gain in energy by removing an interstitial metal atom from the silicon matrix with the interfacial energy cost of enlarging a metal precipitate, one can define a critical radius for a metal precipitate such that for $r > r_{crit}$, a precipitate can grow, but for $r < r_{crit}$ it shrinks.

$$r_{crit} = \frac{2\sigma}{\left(1 + \frac{x}{y}\right)T_{M_xSi_y}kT\ln\left(\frac{C_{M_{int}}}{C^{eq}_{M_{int}}}\right)\left(\frac{C^{eq}_{Si_{int}}}{C_{Si_{int}}}\right)^{\Gamma}}$$

with σ the interfacial energy of the metal precipitate. When $\Gamma > 0$, the injection of silicon interstitials will increase r_{crit}, thereby causing more of the precipitates to shrink. When

$$\left(\frac{C_{Si_{int}}}{C^{eq}_{Si_{int}}}\right)^{\Gamma} \geq \left(\frac{C_{m_{int}}}{C^{eq}_{m_{int}}}\right)$$

the silicide precipitate must shrink, because the critical radius goes to infinity at this point. Since Γ is defined in terms of the volume densities of silicon and the metal silicides, it can be determined from literature

Material	$T_{M_xSi_y}$ (atoms of Si/cm³)	Γ
CuSi	2.14×10^{22}	1.34
Fe_3Si	2.21×10^{22}	0.75
FeSi	4.42×10^{22}	1.13
$FeSi_2$	5.32×10^{22}	−0.12
Ni_3Si	2.32×10^{22}	0.72
Ni_2Si	3.06×10^{22}	0.82
NiSi	4.11×10^{22}	1.22
$NiSi_2$	5.08×10^{22}	−0.03

Table 1 : Volumetric Densities and Γ factors for Common Metal Precipitates

data which silicide precipitates can and cannot be made to shrink. Data for *Cu*, *Fe*, , and *Ni* is shown in Table 1 with the *Cu* data from Das[10] and the rest from Murarka[12].

It is immediately apparent that most of the silicides have $\Gamma > 0$ so that, in the presence of excess silicon interstitials, precipitated metal silicides should dissolve. There are exceptions. Specifically, $FeSi_2$, and $NiSi_2$ have $\Gamma < 0$. As they grow they tend to absorb silicon interstitials. As a result, an excess of silicon interstitials will make these phases grow. Both iron and nickel do decorate stacking faults and, therefore, there is concern over the gettering of $FeSi_2$ and $NiSi_2$. If either of these silicides decorates stacking faults, they cannot be gettered by silicon interstitial injection. Fortunately, $NiSi_2$ does not appear to be one of the phases that ordinarily decorates stacking faults[11]. Interestingly enough, in regions where the concentration of silicon interstitials is high (for example near an oxygen precipitate), nickel does precipitate out as $NiSi_2$. Ourmazd and Schröter [13, 14]. studied the gettering of nickel to a heavily doped phosphorus layer at 900C and to oxygen precipitates during a 752C anneal. In both, the gettered nickel precipitated out and formed separate regions of

nickel silicide and, with TEM, the phase of the nickel silicide was identified as $NiSi_2$ for both. They also studied the gettering of iron to heavily doped phosphorus layers and found that iron precipitated out as $FeSi_2$. Similar behavior was seen by Cullis and Katz[15] looking at iron precipitation in boron diffused junctions. This result implies that boron diffusion acts as a source of silicon interstitials, which has been independently proposed by Morehead and Lever[16] to explain the small tail observed in boron diffusion. These are the only reports of $FeSi_2$ as a stable phase in silicon at any temperature. It seems clear that excess silicon interstitials are causing these phases to preferentially grow. Thus we see that the concept of gettering by silicon interstitials provides a useful framework for understanding the gettering of metal precipitates. For most metal silicide precipitates, an excess of silicon interstitials will help to dissolve the precipitate. There are a few silicides that do not shrink under silicon interstitial injection. For iron and nickel, these forms of the silicide were found to precipitate out in the gettering regions.

Conclusions

In this paper the role of silicon interstitials in the gettering of metal precipitates has been discussed. It was seen that for the most common metallic contaminants, namely copper, iron, and nickel, an excess of silicon interstitials causes most precipitates of these metals to dissolve and thus aids in their gettering. The few metal precipitates that grow larger in the presence of excess silicon interstitials are found only in gettering region which are injecting excess silicon interstitials. It is clear that an effective gettering technique must inject silicon interstitials. Argon implantation, oxygen precipitation, mechanical damage, and phosphorus diffusion all serve as sources of silicon interstitials and thus should be effective in gettering metal precipitates.

References

[1] G.B. Bronner and J.D. Plummer, *J. Appl. Phys.* 61 (1987) 5286.

[2] G.B. Bronner and J.D. Plummer, *Digest of Technical Papers, 1985 Symposium on VLSI Technology* (1985) 18.

[3] D. Lecrosnier, J. Paugam, G. Pelous, F. Richou, and M. Salvi, *J. Appl. Phys.* 52 (1981) 5090.

[4] R. Falster, *Appl. Phys. Lett.* 46 (1985) 737.

[5] L. Baldi, G. Cerofolini, and G. Ferla, *J. Electrochem. Soc.* 127 (1980) 164.

[6] P.F. Schmidt and C.W. Pearce, *J. Electrochem. Soc.* 128 (1981) 630.

[7] A. Goetzberger and W. Shockley, *J. Appl. Phys.* 31 (1960) 1821.

[8] H.H. Busta and H.A. Waggener, *J. Electrochem. Soc.* 124 (1977) 1424.

[9] P.S.D. Lin, R.B. Marcus, and T.T. Sheng, *J. Electrochem. Soc.* 130 (1983) 1878.

[10] G. Das, *J. Appl. Phys.* 44 (1973) 4459.

[11] D.M. Vanderwalker, *Phys. Stat. Sol.(a)* 86 (1984) 507.

[12] S.P. Murarka, Silicides for VLSI Applications, (Academic Press, Orlando, Florida, 1983) 47.

[13] A. Ourmazd and W. Schröter, *Appl. Phys. Lett.* 45 (1984) 781.

[14] A. Ourmazd and W. Schröter, in: *Impurity Diffusion and Gettering in Silicon*, ed. R.B. Fair, C.W. Pearce, and J. Washburn (Materials Research Society, Pittsburgh, Penn., 1985) 25-30.

[15] A.G. Cullis and L.E. Katz, *Phil. Mag.* 30 (1974) 1419.

[16] F.F. Morehead and R.F. Lever, *Appl. Phys. Lett.* 48 (1986) 151.

DEEP STATES IN RAPID ANNEALED SILICON

M. DI MARCO, A.R. PEAKER, C. HILL*, M. HART+ & A.E. GLACCUM#

Department of Electrical Engineering and Electronics, Centre for Electronic Materials
University of Manchester, Institute of Science and Technology
P.O. Box 88, Manchester M60 1QD. England

Deep state measurements have been made on n and p-type silicon grown by Czochralski and float-zone techniques. Major changes are observed when unimplanted slices are annealed. There are differences between lamp and electron beam annealing and between float-zone and Czochralski material. The deep states created in unimplanted layers are present in low concentrations and are associated with vacancy complexes.

1 INTRODUCTION

In principle rapid thermal annealing is an attractive technology for producing shallow junctions and ultra small devices. The initial practical problems encountered (extended defects and enhanced diffusion) are beginning to be understood and expectations are that these difficulties will be overcome in the near future.

However, the third major problem has received rather less attention. When ion implanted material is subjected to rapid thermal annealing (RTA) a high concentration of electrically active point defects is left behind. This can have quite dramatic effects on the device performance. An obvious manifestation is the increase in leakage currents at junctions, but perhaps most significant is a dramatic decrease in both the generation and recombination lifetimes of the material.

Only a few previous publications on this subject exist and most results have been on boron doped Czochralski material [1-3]. In this paper we examine the effects of RTA on as grown material i.e. without implantation. We also draw a comparison between float-zone and Czochralski and between scanning electron beam annealing and much longer anneals using incoherent light. Detailed results on n-type slices are reported in this paper. In addition, our p-type results are compared with previous publications.

2 EXPERIMENTAL

The work was done on 3" float-zone and 4" Czochralski silicon in the resistivity range 20 - 50 Ωcm. The n-type material was phosphorus doped and the p-type boron doped. The oxygen content in the Czochralski and float zone material was $10^{18}cm^{-3}$ and $2x10^{17}cm^{-3}$ respectively (DIN) and the carbon concentration was below $8x10^{15}cm^{-3}$. The electron beam anneal system was used to heat the slices from i) the front face ii) the back face. Anneal times of 1 and 10 seconds at 1100°C were used. A lamp system was employed for longer annealing times and data are reported for 100 seconds at 1100°C anneal. In both cases the initial cool rate was about 100°C s^{-1} but the total cooling time in the electron beam equipment was about half that in the lamp anneal system.

Ohmic contacts were made on the back side of the samples in both the n and p-type cases. Schottky diodes with an area of $0.5mm^2$ were evaporated on the surface (aluminium was used for the p-type and gold for the n-type).

* Plessey (Caswell) Ltd., Towcester, Northants.UK
+ University of Southampton. Highfield, Southampton. UK
# British Telecom Research Laboratories, Martlesham Heath, Ipswitch UK

DLTS measurements were done using a Bio-Rad Polaron DL4600 system and the spectra were recorded at an emission rate of $200s^{-1}$ with -3V reverse bias and a filling pulse duration of 2ms.

3 RESULTS

Figure 1 shows a comparison of DLTS traces from unimplanted n-type Czochralski silicon with and without lamp annealing.

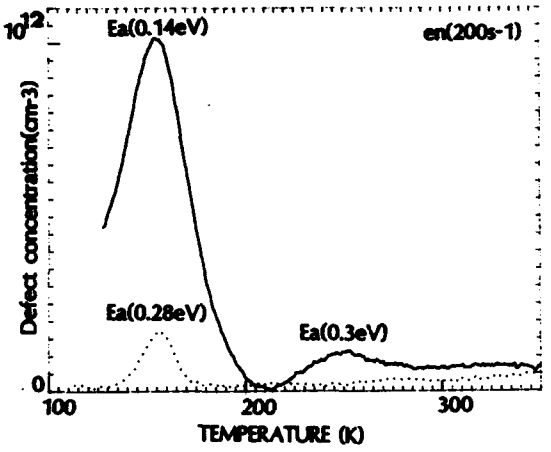


FIGURE 1

Czochralski n-type silicon DLTS spectra recorded at emission rates $200s^{-1}$ for as grown (-). Lamp annealed (...) at 1100°C for 100s

Electron traps with activation energies of 0.14eV and 0.3eV were detected in the unannealed material at concentrations in the range 10^{11} to $10^{12}cm^{-3}$. These electron traps are commonly seen in Czochralski silicon. When the slice is subjected to a lamp anneal both these traps are reduced in concentration by at least an order of magnitude and a new state is apparent with an activation energy of 0.28eV and a concentration of $\sim 10^{11} cm^{-3}$ Although this state has a peak temperature at $e_n = 200s^{-1}$ similar to the 0.14eV state, it is distinctly different - not only in its activation energy but also in its directly measured cross-sections.

The 0.28eV state has similar characteristics to a very commonly observed defect first reported by Yau and Sah [4]. It is associated with a vacancy complex, but no definitive assignment of its physical structure has been achieved.

The most significant point to emerge from these measurements is that there is a dramatic change in the deep level population, resulting from the lamp annealing cycle.

If the same unimplanted material is annealed using electron beam techniques rather different results are observed; these are summarised in Figure 2. In all cases the anneals are at the same temperature as the lamp system (1100°C) and it can be seen that the 0.14 and 0.3eV states present in the starting material have been removed whatever anneal mode has been used. However, new states have been created and these depend on the details of the anneal process. The longest back-face electron beam anneal (10s) creates a 0.26eV state very similar to the vacancy related defect observed in lamp annealed material. It should be noted that although the activation energies are not

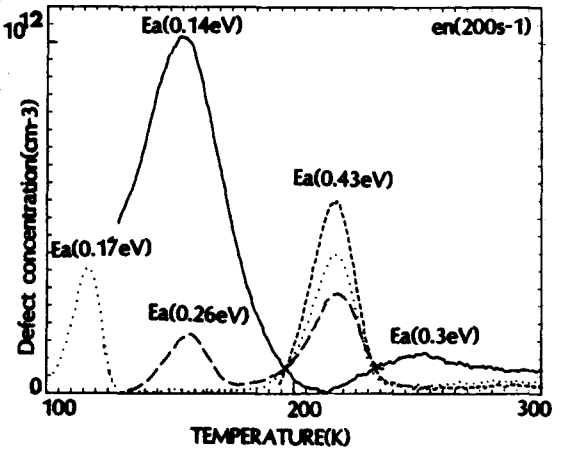


FIGURE 2

Czochralski n-type silicon DLTS results for as grown (-). Electron beam annealed at 1100°C for 1s, back face (...). Electron beam annealed at 1100°C for 10s, back face (- - -). Electron beam annealed at 1100°C for 1s, surface (---)

identical, it is known from previous work that this state has a spread of emission characteristics possibly representing minor variants of the complex. The shorter anneal times (~ 1s) do not produce this state, presumably because of the time needed to generate the complex in the measured region (~ 1 μm from the surface). However, in all the electron beam anneal cases a state with an activation energy of 0.43eV is generated, its concentration varying as shown.

This state has characteristics identical with those ascribed to a vacancy-phosphorous associate while the state at 0.17eV is a vacancy-oxygen complex [5-6]. A recent paper by Song et al [7] expresses the view that the long standing assignment of the 0.43eV state as the vacancy-phosphorus pair is somewhat simplistic. The basis for this is that they have observed defect reaction occurring at low temperature (<30°C) under specific bias conditions which convert the 0.43eV state into other defects with very different emission properties. These reactions are reversible. Although we have observed similar multistable complexes and, in addition, metastable configurations of defects in silicon [8] these do not occur in these samples. Indeed we observe multistability and metastability to be common in annealed silicon which has been damaged by implantation or irradiation but not in normal as grown material subjected to RTA. It is our belief that the reason for this is the relatively small concentration of defects present in as grown material. Both the 0.43 and the 0.17eV states are commonly observed in electron irradiated silicon, but anneal out at low temperatures (<300°C) when using conventional furnace treatment. Their presence and that of the other state at 0.28eV in the lamp annealed material is conclusive evidence of vacancy reactions occurring during the anneal period and the cool-down cycle. Work is now in progress to study the reaction pathways in detail but it is evident that the anneal time is a critical parameter.

In the case of the float-zone material, the initial population of deep states is completely different to the Czochralski slices. This is shown in Figure 3. The state observed in the highest concentration (0.56eV) has an emission characteristic which cannot be associated with a specific defect as it coincides with that observed for a range of chemical species and intrinsic states [9].

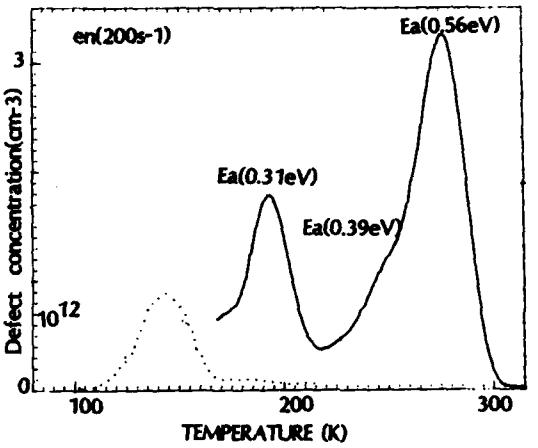


FIGURE 3
Float zone n-type silicon DLTS spectra for as grown (-). Lamp annealed at 1100°C for 100s (...)

The other deep levels have not been linked to previously published results. After RTA all these states disappear and the 0.28eV vacancy complex, seen in the annealed Czochralski material, is present in a concentration of ~ 10^{12} cm $^{-3}$ (an order higher than in the Czochralski material). There was no evidence of the vacancy-oxygen complex at 0.17eV seen in the Czochralski material. This was perhaps predictable because of the much lower oxygen concentration in the float zone material.

In the case of p-type layers, our results are not as consistent as those obtained using n-type slices. However, in general they are similar to those of Barbier et al [3]. He observes the annhilation of all deep states for lamp anneals longer than 20s.

But in our case, a hole trap with an activation energy of 0.45eV is present after a 100s lamp anneal in a concentration in the range $2-9 \times 10^{11} cm^{-3}$. The emission characteristics for holes are described by:

$$e_p = 7.17 \times 10^6 T^2 \exp(-0.45/kT)$$

This same state is observed by Barbier in quite high concentrations (up to 10^{14} cm^{-3}) for anneals of a few seconds at 1000°C. Borenstein [2] links a similar state to interstitial iron.

4 CONCLUSIONS

In annealed material which has not been implanted perhaps the most important feature to note is that the deep state concentrations are low, in most cases lower than the as grown slices. Indeed Shockley-Hall-Read calculations indicate that it is unlikely that the defect population produced by the lamp anneals could have any significant device implications either in MOS or bipolar technology.

However, this must be regarded as a tentative conclusion as additional measurements need to be made on the minority carrier capture cross-section of some of the states before a definitive pronouncement can be made on the effect they may have on the low level recombination lifetime. It is obvious from these results that the defect reaction occurring during RTA involve vacancies, the shallow dopant species and inadvertent impurities in the crystal.

In the case of implanted material, the concentration of intrinsic defects in the early stages of annealing is vastly greater. Bearing in mind the origin of the electrically active deep states observed in this work, it is not surprising that the deep level concentrations detected in material which has been ion-implanted and rapid annealed are several orders higher than in the slices measured here. It is for this reason that they are of such importance in devices.

Preliminary work indicates that there are marked similarities in the defect reactions occurring in unimplanted and implanted material during rapid thermal annealing, although the higher concentration of defects in implanted layers seem to allow the formation of a number of alternative configurations of defects related to those observed here.

ACKNOWLEGMENTS

We should like to thank SERC and British Telecom for providing financial support for this work. Our thanks are also due to B. Hamilton, J.T. Thornton, A. Evans, J.R. Davies and N.O. Pearce for helpful discussions and J. Barton, M. Harding and C. Graham for assistance with the preparation of samples and the manuscript.

REFERENCES

[1] Pensl, G., Schulz, M., Stolz, P., *Energy Beam-Solid Interactions and Transient Thermal Processing*, Mat.Res.Soc.Proc. (North-Holland, New York, 1984), p 347

[2] Borenstein, J.T., Jones, J.T., Corbett, J.W., Oehrlein, G.S., Kleinhenz, R.L., Appl.Phys.Lett.49 (1986) 199

[3] Barbier, D., Remram, M., Joly, J.F., Laugier, A., J. Appl.Phys. 61 (1987) 156

[4] Yau, LD. and Sah, C.T., Solid State Electron. 14,(1971) 193

[5] Kimerling, L.C., *Radiation Effects in Semiconductors*. Inst.Phys.Conf.Ser. No.31, (1976) 221

[6] Song, L.W., Benson, B.W. and Watkins, G.D., Phys.Rev.B. 33 (1986) 1452

[7] Walker, J.W. and Sah, C.T., Phys.Rev.B. 7 (1973) 4587

[8] Hamilton, B., Peaker, A.R. and Pantelides, S.T., *Entropy Driven Metastability of Deep Levels*. To be published 1987

[9] Lang, D.V., Grimmeiss, H.G., Meijer, E. and Jaros, M., Phys.Rev. B22 (1980) 3917

GETTERING AND DEEP STATES IN P-TYPE CZOCHRALSKI SILICON

N. JHA *; A.R. PEAKER & G. KEEFE-FRAUNDORF+

Department of Electrical Engineering and Electronics, Centre for Electronic Materials, University of Manchester, Institute of Science and Technology
P.O. Box 88, Manchester M60 1QD, England

This paper reports deep state measurements on boron doped Czochralski silicon. A comparison is made between four gettering technologies after the slices have been subjected to oxidation cycles in steam at 1100°C for two hours. An analysis of the results is made using a diffusion model of gettering.

1 INTRODUCTION

An important part of device fabrication is the use of gettering regions remote from the active junctions. These regions accumulate unwanted impurities introduced during processing. In most systematic studies of impurity distribution after gettering slices have been contaminated with easily detectable impurities, such as gold, and the reduction in their concentration observed after appropriate heat treatment. The introduction of high concentrations of an impurity is necessary because existing chemical techniques of analysis are not sufficiently sensitive to study inadvertent impurities (i.e. those impurities introduced during processing or already present accidentally in the slice). There are two major problems with the contamination approach, firstly that the gettering mechanisms may be concentration dependent and secondly that the effect on the intentionally introduced impurity may be significantly different to that on those species present inadvertently. In either case, the results will not be representative of normal processing. An alternative but less systematic approach has been to fabricate devices using different gettering techniques. Here the hold time of MOS capacitors or the gain of a bipolar transistor has often been used as a measure of the gettering efficiency. The fundamental disadvantage of this experimental method is that many other factors can also affect the measured parameters.

2 EXPERIMENTAL

In the work reported here we have studied heat treated slices gettered in different ways and compared them with the starting material. The slices were (100) and (111) Czochralski silicon boron doped with resistivities in the range 5 - 20 Ω cm. The oxygen content was $\sim 10^{18} cm^{-3}$ with carbon $< 10^{16} cm^{-3}$. One group of wafers had the back surface damaged mechanically by abrasion with 12 μm particles (MD). Another group was laser scanned on the back face using a 12 watt niobium-YAG laser which produced spots of damage 20 μm in diameter and 40 μm apart (LD). The third group was phosphorus diffused (PD) and the fourth group is referred to as enhanced gettered (EG). The measurements on the EG slices have been undertaken in the surface layer denuded of oxygen which in this case was 15 μm thick. The EG slices had been heat treated to precipitate oxygen in the bulk of the crystal and a polycrystalline layer deposited on

* B.P. Research Centre, Sunbury on Thames. U.K
+ Monsanto Electronics, St. Louis, Missouri. U.S.A.

the back face. All the slices were oxidised in steam at 1100°C for two hours and then the front face oxide removed. This was repeated up to four times. After each oxidation samples were extracted and the oxide removed. Schottky barriers were then evaporated onto the front surface, except for the phosphorus gettered case where the phosphorus diffused region was etched off before the barriers were evaporated. Deep Level Transient Spectroscopy (DLTS) measurements were made on all the layers. Some slices were processed to produce MOS capacitors with an 80 nm gate oxide grown at 1000°C followed by an anneal cycle. These were then used for generation lifetime measurements using the Zerbst technique. A disc was cut from each slice and thinned to examine the backface using transmission electron miscroscopy (TEM). The number and size of dislocations and stacking faults were measured in order to compare the gettering sites in the samples. Measurements were also made of the front face etch pit density using a Sirtl etch for 4min.

3 RESULTS

Two majority carrier traps were observed in all the samples in concentrations in the range 10^{11}–$10^{15}cm^{-3}$. The hole emission characteristics of these states are described by:

$$e_p = 8.7 \times 10^5 \, T^2 \exp(-0.376/kT)$$

and:

$$e_p = 4.9 \times 10^3 \, T^2 \exp(-0.428/kT)$$

The directly measured hole capture sections of these states is $\sim 10^{-19}cm^2$ with a much larger electron cross section ($\sim 10^{-15}cm^2$). Estimates of the recombination generation rate via the centres using Schockley–Hall–Read kinetics indicate that both states could degrade the low level minority carrier lifetime in p-type material in the detected concentration and that the 0.43eV state could be a significant generation centre likely to affect the performance of MOS devices. However, no positive identification of the impurity or defect species responsible for the deep states has been made although the behaviour is suggestive of complexes associated with the 3d transition metals.

Figure 1 reports the concentrations of the two deep states after three oxidation cycles in slices using the different gettering technologies. The control is the starting material. It is apparent that there are very dramatic differences between the efficiences of the various gettering techniques. Also, rather interestingly, the effectiveness of a particular gettering method can be very different for each of the two traps studied. For both traps the phosphorus diffusion is the most effective, the next best being enhanced gettering. It must, however, be remembered that the mechanical damage technique and the laser damage method are entirely back face processes. The enhanced gettering method is a mixture of back face and bulk gettering while phosphorus diffusion gettering actually takes place within a few microns of the active region. Even apart from this complication, the comparison is not straightforward essentially because the gettering efficiency changes with the number of oxidation cycles. This is shown in the laser damage case in Figure 2.

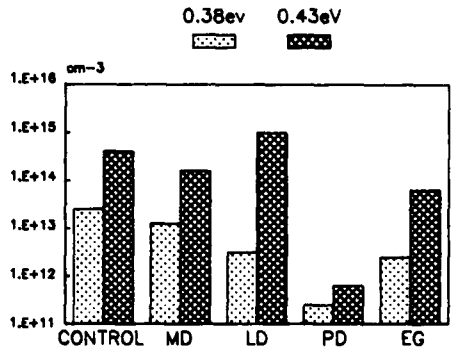


FIGURE 1

The concentration of the 0.38 and 0.43eV deep states after three oxidation cycles in slices using different gettering techniques.

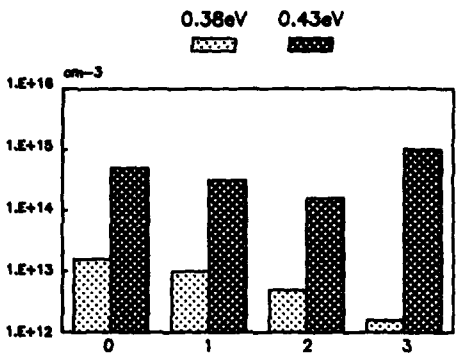


FIGURE 2

The concentration of the two deep states in material with back face laser damage after one, two and three heat treatment cycles at 1100°C for two hours in steam

It can be seen that the 0.38eV state is consistently reduced in concentration. However the 0.43eV level reduces initially and then increases. It appears that the higher concentration state is not only present in the starting material, but is also being introduced during processing. The heat treatments took place in normal production line furnaces and there is no reason to believe they were particularly contaminated. The gettering process for this state in laser damaged material undergoes a saturation whereas it does not for the lower concentration state at 0.38eV. This observation has very significant implications in elucidating a gettering mechanism.

Saturation does not occur in the case of phosphorus diffusion gettering and is not particularly significant in the enhanced gettered material.

The TEM measurements give some indication of the number of gettering sites available on the back face. Table 1 compares the number of defects observed after three oxidations. In this case they were made up entirely of stacking faults, the dislocation density was insignificant. The last column which indicates the total gettering length density, provides a good comparative figure.

It is interesting to note that the enhanced gettered material has over a hundred times more gettering sites than the mechanical damaged layers. One might reasonably expect the longevity of the enhanced gettered slices to be considerably greater than those mechanically damaged or laser damaged and, indeed, this is observed to be the case.

TABLE 1

Back Face Defects After Three Oxidations

Type of Gettering	Size μm	Density cm^{-3}	Total Gettering Length Density $cm/cm^3(cm^{-2})$
MD	19	$3x10^8$	$5.7x10^8$
LD	26	$8x10^8$	$2.1x10^7$
EG	0.4	$2x10^{13}$	$8\ x10^8$

However, it is important to appreciate that the number of defects available changes very significantly with the number of oxidation cycles. Table 2 gives an indication of these changes in laser damaged material. It can be seen that both the size and density of the defects changes and the net effect is to increase the number of gettering sites between the first and third oxidations by an order of magnitude. They then start to reduce. This provides a qualitative explanation of the saturation observed for the higher concentration trap in *Figure 2.*

TABLE 2

Back Face Defects in Laser Damaged Material

No. of Oxidations	Stacking Faults Size μm	Density cm^{-3}	Total Gettering Length $cm/cm^3(cm^{-2})$
1	10	$2x10^8$	$2.1x10^6$
2	23	$2x10^8$	$4.6x10^6$
3	26	$8x10^8$	$2.1x10^7$
4	30	$2x10^8$	$6.3x10^8$

Numerous models of gettering have been proposed in the literature. Vengurlekar [1] has derived a relatively simple but convincing relationship for the residual concentration of impurities in a gettered layer based on the diffusion of the contaminent species to a region of defined thickness. Using this model for the 0.38eV trap we have one variable parameter, the diffusivity in the gettered layer. To fit our results the diffusivity must be in the range 10^{-7} to $10^{-8}cm^2s^{-1}$ a value which is reasonably sensible for transition metal diffusion.

However, if we analyse the enhanced gettered material perhaps the most interesting point to emerge is that after the first oxidation cycle the gettering takes place almost exclusively in the poly-silicon layer. If we assume that the dominant mechanism is intrinsic gettering via the oxygen precipitates then we require the diffusivity for the species which diffused at $10^{-7}cm^2s^{-1}$ in the MD layer to be $10^{-10}cm^2s^{-1}$ in the EG material.

Consequently, our model of the processes occurring in the enhanced gettered slices is that during the first oxidation cycle the intrinsic precipitates play a significant role, but subsequently the gettering is dominated by the back face. This is true, even though additional precipitates appear to be formed during the thermal cycles as judged from measuring the interstitial oxygen concentration using optical absorption.

The dominant question, however, is what effect does this have on the electrical properties. We predict from the deep level measurements that the enhanced gettered material should be considerably better. Table 3 lists the generation lifetimes measured using the Zerbst technique.

TABLE 3
MOS Lifetime After Three Gettering Cycle

Gettering	Generation Lifetime
None	0.5-1.3μs
MD	0.5-1.5μs
LD	0.8- 2μs
EG	2 - 5μs

There is little doubt that among the slices measured the enhanced gettered material is vastly superior. A comparision of the etch pit density on the front surface supports this view. The etch pit density after four oxidations is $10^5/cm^2$ on the mechanically damaged wafers, where it is less than $300/cm^2$ in the enhanced gettered material.

4 CONCLUSIONS

The DLTS measurements have shown that there are vast differences between the effectiveness of various gettering techniques, particularly in terms of their longevity. TEM measurements indicate that a good reason to explain this behaviour is simply that there are very large differences in the number of gettering sites available in the different types of material. Measurements of generation lifetime on MOS capacitors indicate that the deep level measurements are a good indicator of device performance and confirm the view that enhanced gettering provides a number of advantages over traditional back faced damage techniques.

REFERENCES

[1] Ver[illegible]r, A.S., Appld.Phys.Lett. 41 (3)0

Session A2.4

CMOS Technology

Chairman: L. Baldi

Tuesday, September 15, 1987

CMOS TECHNOLOGY WITH SELF-ALIGNED CONTACTS AND SELF-ALIGNED SILICIDE*

J.-M. MORET, P. WEISS, H. LUGINBUEHL

Swiss Center for Electronics and Microtechnology Inc.
Maladière 71, CH-2007 Neuchâtel, Switzerland

M. DUTOIT

Swiss Federal Institute of Technology
CH-1015 Lausanne, Switzerland

Simple process modifications are proposed to notably increase the packing density of a given CMOS technology. These include the self-alignment of source-drain contacts together with the self-aligned silicidation of the diffusion regions and the polysilicon lines. The required gate sealing makes use of the sidewall spacer technique. The proposed changes have been incorporated into our conventional CMOS technology. The use of self-aligned contacts and self-aligned silicide allows to reduce the circuit surface of a C^2-MOS latch by a factor of 2.4 and the gate delay of a ring oscillator by a factor of 1.6.

1. INTRODUCTION

Structural improvements of conventional CMOS devices can provide a large increase in packing density, while keeping the same minimum dimensions. This approach is particularly attractive for gate lengths in the 2 µm range because it may avoid further down-scaling, which would require specific efforts to overcome parasitic effects due to short channels and high electrical field. For submicron dimensions, resorting to self-aligned techniques will become essential since alignment tolerances won't follow the size scaling any longer.

Several techniques have been proposed for source-drain contacts self-alignment [1 to 5]. The widely used principle is to seal the gate with an extra dielectric layer so that the source-drain contact can be etched without risking any electrical short between the electrodes, even if the contact hole overlaps the gate. Gate sealing generally implies polysilicon gate oxidation, a process which proves unstable for short devices. To our knowledge the process proposed by J. Solo de Zaldivar [4] is the only one used today at an industrial production level [6].

The purpose of this work was to take full advantage of the structural improvements obtained by combining the self-aligned source-drain contacts with self-aligned silicide [7]. Special attention has been paid to the wiring of the MOSFET with its surrounding interconnect lines, resulting in a consistent and very efficient set of layout rules. The technique we propose only resorts to well established low-temperature steps, thus lending itself to its incorporation into a CMOS VLSI process.

2. DEVICE STRUCTURE

The MOSFET structure we have realized is shown in fig. 1 and 2. Source and drain areas are minimized. Polysilicon interconnect lines and diffusion regions directly adjacent to them can either be connected to or insulated from each other.

Work sponsored by the Swiss National Foundation for Scientific Research (PN 13).

The self-aligned silicide layer is formed on diffusion regions and polysilicon lines, excluding the gate area. This avoids the back-end process problems usually encountered in standard salicide technology [8]. The price to pay for the partial silicidation is obviously a slight increase of the overall line resistance.

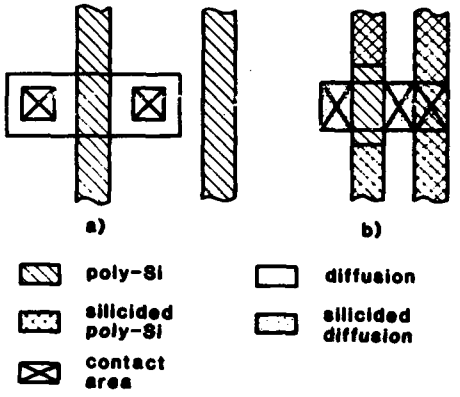


FIGURE 1

Comparison of MOSFET and interconnect line layout. a) conventional b) with self-aligned contacts and self-aligned silicide (interconnect line connected to the drain).

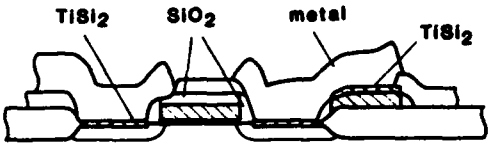


FIGURE 2

Cross section of the structure given in fig. 1b.

The effectiveness of the resulting layout is illustrated in fig. 3, using a C^2-MOS latch circuit as an example. The circuit surface has been reduced by a factor of 2.4 relatively to the conventional CMOS process. The circuit has been synthesized with an automatic program based on the concept of "gate matrix" [9].

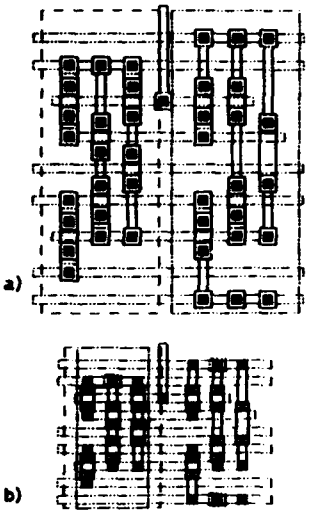


FIGURE 3

Comparison of C^2-MOS latch circuit
a) conventional
b) with self-aligned contacts

3. FABRICATION PROCESS

Up to polysilicon layer deposition and doping a standard 2 µm CMOS process is followed. Self-aligned source-drain contacts require both top and side sealing of the polysilicon gate. The top insulator consists of a 300 nm thick deposited SiO_2 layer. The oxide is patterned in two steps. First it is removed by wet etching in all areas where surface insulation of the polysilicon layer is not necessary (fig. 4a). This requires a non-critical photolithographic step, using an extra mask. It should be noted that such a supplementary step is characteristic of any self-aligned source-drain contacts scheme. Then the polysilicon lines, with their patterned top oxide layer, are defined by a two-step reactive plasma etch (fig. 4b).

The use of successive highly selective and anisotropic etch processes allows to cope simultaneously with the covered/non covered polysilicon regions, leaving vertical sidewalls. The lateral gate insulation makes use of the well established sidewall spacer technology [10] (fig. 4c). Two kinds of SiO_2 spacers are formed, depending if the polysilicon line is insulated or not on the top. Their heights are very different (0.48 µm and 0.25 µm respectively) but their widths proves quite similar (0.20 µm and 0.22 µm respectively).

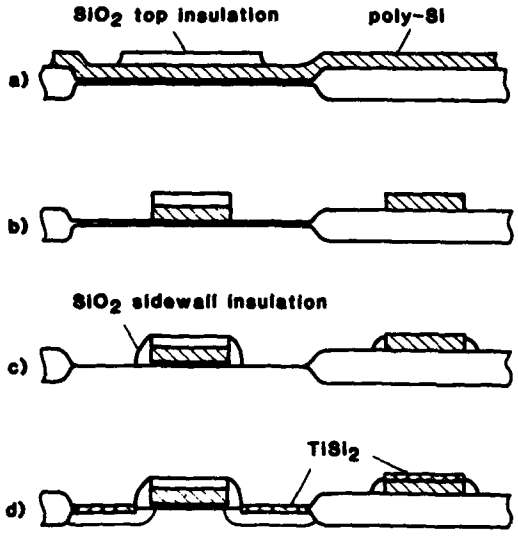


FIGURE 4

Process sequence to form the gate insulation and the self-aligned silicide on source-drain regions and on polysilicon lines (outside gate areas).

After p^+ and n^+ source-drain implantation and drive-in, titanium silicide is formed on diffusion regions and on non-covered polysilicon lines using the salicide technique (fig. 4d). For this purpose, a 40 nm thick titanium layer is sputtered onto the wafer. The silicidation is then accomplished in two steps, resorting to a rapid thermal processing equipement. Diffusion and polysilicon sheet resistances are both lowered to 2-4 Ω/□.

At this point of the process a 200 nm thick phosphosilicate glass layer is deposited and conventional processing resumes.

4. RESULTS

Fig. 5 shows a SEM cross section of the realized gate structure after contact window opening. The gate encapsulation has been partly etched during the later step but its thickness is still sufficient to leave the gate to drain leakeage unchanged (fig. 6). The breakdown voltage exceeds 25 V.

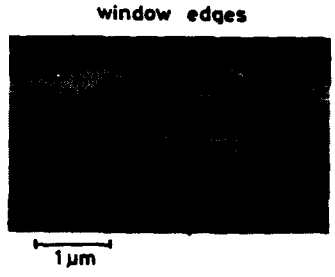


FIGURE 5

SEM cross section of a gate structure after contact window etching.

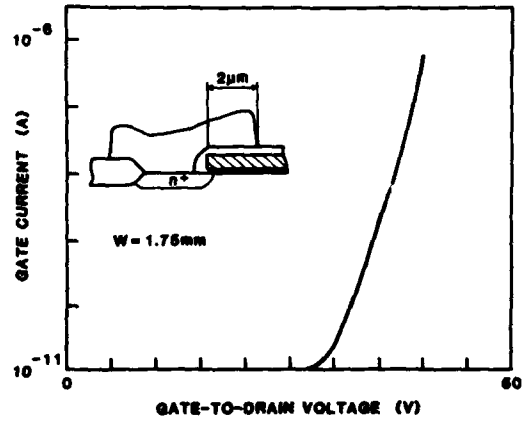


FIGURE 6

Typical gate to drain leakage characteristic (contact window overlaps gate region).

The capacitance properties of the self-aligned contacts MOSFET differ from those of the conventional one in two points : i) the drain junction capacitance is lowered by a factor of ~ 2 because of the drain surface reduction, ii) at zero overlap, the metal contacting the drain significantly contributes to the gate-to-drain capacitance; but, due to the oxide spacer, the usual gate-to-drain overlap is effectively reduced and thus the measured capacitance proves lower than for the conventional structure (fig. 7).

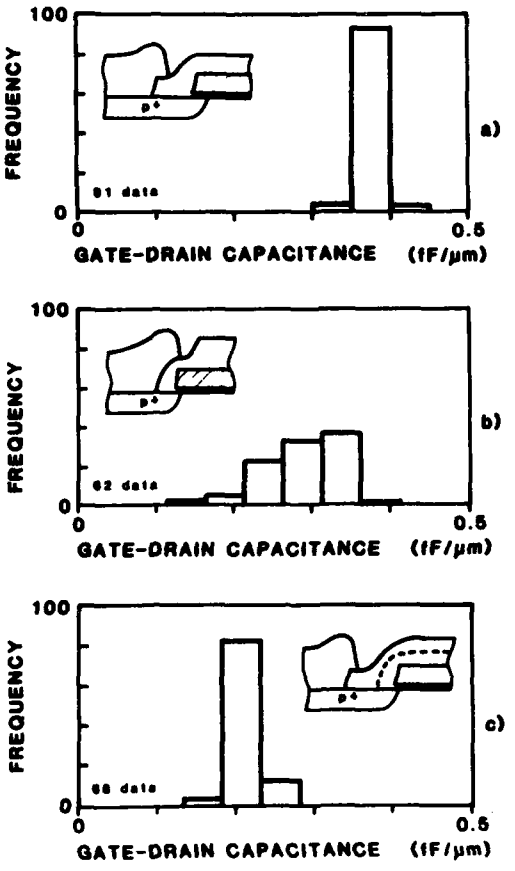


FIGURE 7

Single wafer gate-to-drain capacitance histograms. a) conventional structure b) with self-aligned contacts (zero nominal window, metal and gate overlap) c) like b) but with window and metal 2 μm distant from the gate.

Gate-to-drain capacitance at zero overlap (fig. 7b) are quite scattered; this dispersion is due to the several photolithographic steps and etches involved.

The reverse-biased drain characteristics of MOSFETs with self-aligned contacts are very similar to those of conventional ones. This is also the case for most others parameters, such as threshold voltage, subthreshold slope and transconductance (fig. 8).

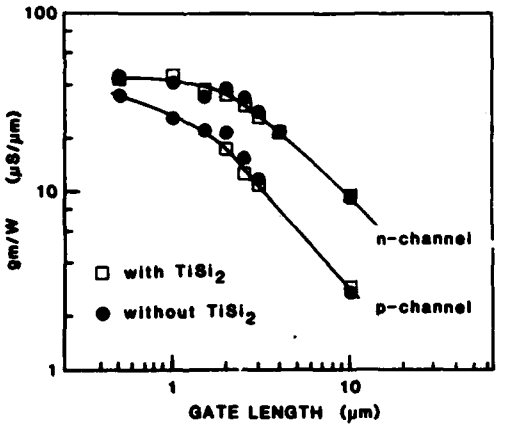


FIGURE 8

Transconductance vs. gate length of MOSFETs with and without $TiSi_2$.

21-stage ring oscillators have been fabricated with the proposed process. The gate delay per stage is reduced, relatively to the standard technology, by a factor of 1.6 from 0.70 ns to 0.44 ns (at 5 V).

5. CONCLUSIONS

By exploiting the well established sidewall spacer process, a self-aligned source-drain contacts technique has been developed on a CMOS technology. The proposed solution further allows self-aligned silicidation of the diffusion regions and of the polysilicon lines outside gate areas. Only low temperature steps are involved.

Characteristics of MOSFETs with self-aligned contacts and titanium silicide have been found very similar to those of conventional ones. Use of self-aligned contacts lowers the drain junction capacitance as well as, in our case, the gate-to-drain capacitance. Self-aligned source-drain contacts with self-aligned silicide together provide a large increase in packing density and circuit speed. Moreover the resulting layout rules allow very regular designs, which prove most convenient for automatic layout generation.

REFERENCES

[1] M. Sakamoto and K. Hamano, IEDM Techn. Dig., (1980) 136.

[2] K. Ohta, K. Yamada, M. Saitoh, K. Shimizu and Y. Tarui, IEEE ED-27 (1980) 1352.

[3] M.K. Khan and G.C. Godejahn, J. Electrochem. Soc., 128, (1981) 1333.

[4] J. Solo de Zaldivar, Proc. NTG Fachtagung, Grossintegration, Baden-Baden, 16-18 März, 1981.

[5] T. Yachi and N. Yamauchi, IEEE ED-29, (1982) 143.

[6] R.E. LUscher and J.S. De Zaldivar, ISSCC Tech. Digest (1985) 260.

[7] C.K. Lau, Y.C. See, D.B. Scott, J.M. Bridges, S.M. Perna and R.D. Davies, IEDM Tech. Dig. (1982) 714.

[8] C.Y. Ting, F.M. d'Heurle, S.S. Iyer and P.M. Fryer, J. Electrochem. Soc., 133 (1986) 2621.

[9] C. Piguet, J. Zahnd, A. Stauffer and M. Bertarionne, IEEE, J. Solid-State Circuits, SC-19, (1984) 425.

[10] S.H. Dhong and E.J. Petrillo, J. Electrochem. Soc., 133 (1986) 389.

1µ MOS DEVICES WITH SELF-ALIGNED TITANIUM SILICIDE AND CVD TUNGSTEN AS FIRST METALLISATION LEVEL

C. Arena, S. Deleonibus, G. Guegan, P. Laporte, F. Martin, J.L. Pelloie
LETI/IRDI - COMMISSARIAT A L'ENERGIE ATOMIQUE CEN/G - 85 X - 38041 GRENOBLE CEDEX - FRANCE

I. INTRODUCTION

Ti Silicided CMOS technology has been extensively studied in the last few years, because scaling of device sizes towards submicronic field needs reduction of source and drain junction depth. The increase of junction sheet resistance can be a problem for keeping optimized device performances The main circuit applications concern poly-Si gates [1] and silicided junctions [2]. The great advantages of self aligned silicidation are the improvement of sheet resistance compared to N+ or p+ layers, the possibility of higher packing density, especially with the introduction of butted contacts and a best immunity to latch-up [2,3,4]. An improvement on speed can be expected for some products as ROM and PLA by drain-source silicidation and for RAM by gate silicidation.

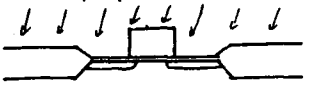
CVD tungsten as first level metallization was chosen for its ability to meet some of the most severe requirements in submicronic multilevel metallizations [5]. Its exellent conformal coverage compared to the problems associated with physical vapor deposition of aluminium allows a complete filling of small contact holes. Morever W CVD shows no hillock formation thus allowing far higher thermal treatments to achieve a good planarizing and reliable interlevel dielectric. Further, the electromigration resistance of W is much greater than that of Al [5].

2. PROCESS DESCRIPTION

Figure 1 shows the main steps of the process flow used on these experiments. Starting with both NMOS and PMOS bulk silicon materials, a conventional LOCOS and a 20 nm gate oxide are achieved. After threshold adjust implantations, 420 nm of polysilicon are deposited. Poly is then doped from a PO Cl_3 source before achieving gate etching. Phosphorus implantation is used for self-aligned LDD on NMOS just after poly delineation through the remaining gate oxide. Spacer technology is then used for LDD NMOS control, PMOS lateral diffusion control as well as poly to source-drain isolation in the $TiSi_2$ silicidation process. 350 nm of PECVD SiO_2 are deposited and etched in an RIE reactor to get the desired spacer length. Prior to n+ and p+ implantations, source and drain are oxidized in order to have an implantation mask on poly and substrate. Once the implantations are done, the structure is furnace annealed at 950° C for activation and diffusion of dopants.

Fig. 1 Silicide and tungsten metallization main steps process.

a.



After gate etch, n-shallow implant

b.



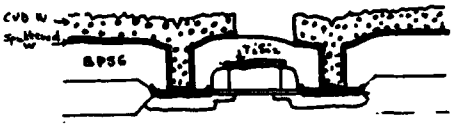
Spacer definition, source & drain oxidation n + & p+ implants, anneal

c.



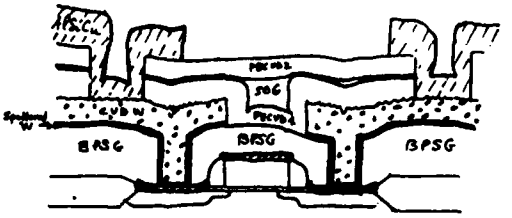
Oxide deglaze, salicide process

d.



BPSG deposition, Reflow, contact mask & etch, Sputtered W deposition, CVD W deposition, Metal 1 mask & etch.

e.



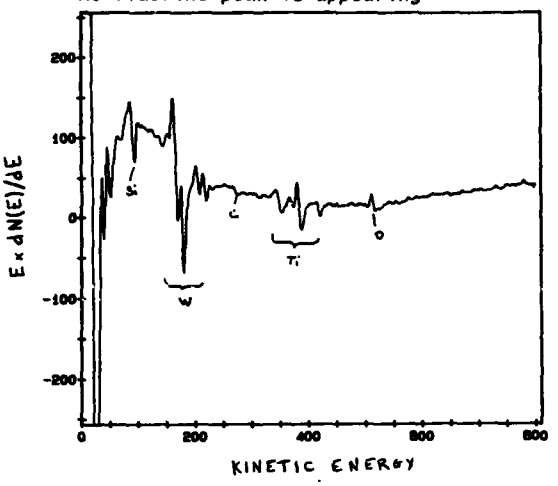
Metal 2 - Metal 1 planarization, via mask & etch, AlSi(1 %) Cu(0.5%) deposition, métal 2 mask & etch.

After stripping of thermal oxide on junctions and poly with BHF 60nm titanium are sputtered and annealed at 625°C under nitogren ambiant in high flow purged furnace. After TiN elimination in perodixe-sulphuric bath, and second heat treatment at 850°C 10s with rapid termal annealing system (AET) under Ar ambiant $TiSi_2$ reaches 2.5Ω/□ on bare silicon wafers without implants. Then BPSG is deposited and reflow is achieved at 850°C under steam ambiant.

Chemical vapor tungsten deposition is achieved in a hot wall reactor (ANICON) at 400°C and a total pressure of 0.4 torr, through the H_2 reduction of WF_6 (one step process) with a ratio $PH_2/PWF_6 \simeq 5$. In these conditions, no uniform and adherent W deposit can be obtained on oxides. A "glue" layer between BPSG and CVD W is necessary. For obvious chemical compatibility, a sputtered W layer is chosen and optimized in terms of good adherence of the sputtered and CVD W bilayer on BPSG and in terms of good electrical contact with the $TiSi_2$ underlayer. Therefore these requirements lead to the following sputtering conditions : first an in situ back sputtered etch of $TiSi_2$, then W deposit in an Alcatel PUMA 500 system at 2000 Watt under an Argon ambiant (P Ar $\simeq 2.10^{-2}$ torr), the substrate is maintained cold (< 100°C).

Another point of importance is the problem of titanium solid fluorine species (TiF_3, TiF_4,) formation due to the possible reactic of $TiSi_2$ with $WF_6 + H_2$ and their byproducts [6]. 150 nm sputtered W is expected to act as an active fluorine gas diffusion barrier in regards to the underlayered $TiSi_2$. The Auger depth profile spectrum reveals no trace of fluorine presence at the exact interface between sputtered W and $TiSi_2$ (fig. 2). Moreover good adherence (scotch tape test) on full wafers is obtained Contrary to direct CVD W deposition on $TiSi_2$. Recently Broadbent [7] reported analogous results and confirmed the role of sputtered W as a diffusion barrier. A sputtered W thickness of 15 nm should be sufficient to prevent from unwanted WF_6 Ti reactions.

Fig. 2 Auger profile of $W/TiSi_2$ interface
No fluorine peak is appearing



At last one interesting feature concerns the double metallization level interdielectric which was obtained through whole uniform 400°C PECVD oxyde deposits leading to a minimal stress level and better mechanical reliability. It can be noted that for an aluminium alloy metallization severe limitation of temperature make the intermetal dielectric realisation difficult due to hillocks formation. Al Si (1 %)-Cu (0,5 %) was deposited and delineated as second level metallization.

At witness process is achieved in the same time with aluminium metallizations (first level) and poly gates for comparison with $TiSi_2$, W technology.

3. TOPOGRAPHIC RESULTS

Figure 3 illustrates the tremendous planarization effect of both BPSG (5% B, 4% P) reflowed at 850 C and CVD W on a periodic structure of poly gates.



Fig.3 - ILLUSTRATION OF BPSG AND TUNGSTEN PLANARIZATION EFFECT

Figure 4 illustrates the small holes filling capacity of CVD W due to an isotropic growth mechanism. Ti Si_2 is revealed in black with an HF chemical etch. The sharp interface between W and Ti Si_2 can be noted.

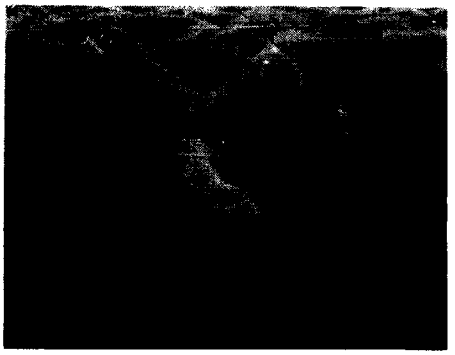


Fig.4 - QUASI TUNGSTEN PLUGGED CONTACT HOLES (1.5μm)

4. ELECTRICAL RESULTS

a. Sheet, contact and via resistances

Higher sheet resistance value on n+ (4Ω/□) and p + (3.5Ω/□) junctions than previously observed on bare silicon results from steam ambiant of reflow and has yet to be optimized.

Specific contact resistivities are measured on Kelvin resistor and found to be 20Ω.μm² for Ti Si_2/n + and 100Ω. μm² for Ti Si_2/p+. These results are in good agreement with others previously reported [8].

Via resistances between W as first metal and AlSiCu as second metal are found to be 80 mΩ for 1.2 μm size vias.

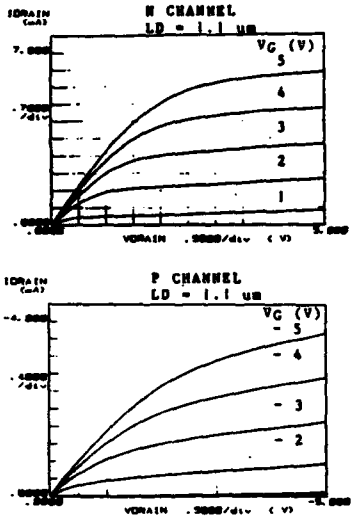
b. Active devices characteristics

Table 1 lists the different parameters obtained for a transistor having an effective length of 0.8 μm (W_{eff} = 19 μm). Figure 5 shows typical I-V characteristics for a drawn lengh of 1.1 μm. These characteristics are quite similar with and without Ti Si_2. All transistors parameters have the same values than those obtained for standard technology (n+ poly gate, Al as first metal and no silicidation).

Table 1. Electrical parameters of n and p channel transistors (L_{eff} = 0.8 μm - W_{eff} = 19 μm).

	Threshold voltage (V)	Subthreshold Slope (mV/dec)	Breakdown voltage (V)	Mobility ($cm^2.V^{-1}.s^{-1}$)
NMOS	0.4	90	7	530
PMOS	- 1.1	95	- 10	180

Fig. 5 : I_{DS} versus V_{DS} characteristics on silicided PMOS and NMOS transistors. No difference with no silicide process transistors.



We have noted a significant improvement on punch-through voltage for NMOS device with silicidation. Figure 6 shows the evolution of punchthrough voltage versus effective length for n-channel transistors with and without Ti Si_2. Punchthrough voltage is defined as the drain voltage associated with a drain current of 10pA per transistor width unit (the maximum value on the curve is voluntarily limited to 8V), the other electrodes being grounded. We think that this effect is due to a better location of carriers injection in the channel as the junctions are silicided. This improvement has not been observed on p-channel transistor which is a buried channel transistor.

Fig. 6 BVDSS versus leff. Comparison is given for silicide and no silicide prrocess.

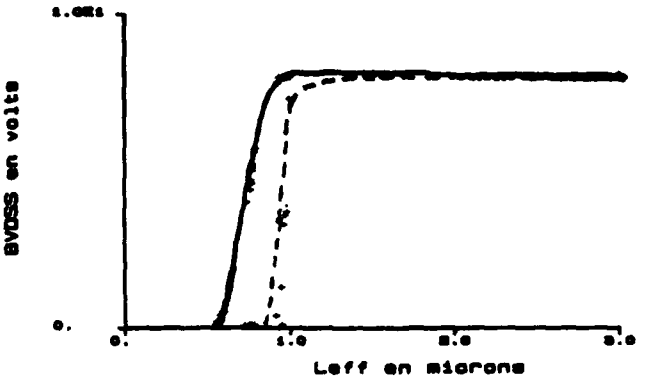


Fig. 6 - NMOS —— silicide process
--- no silicide

5. CONCLUSION

We have demonstrated in this paper that a technology including $TiSi_2$ sell aligned silicidation of junctions and n-poly gates plus W as first level metallization can be performed without any electrical degradation compared to conventional technology. This process will be applied to a future 0.8 µm drawn CMOS technology.

The authors thank the whole people of R & D team who contributed to the study.

REFERENCES

[1] Y. Murao et al., IEDM Tech. Dig., p. 518, 1983.

[2] Y. Taur et al., IEEE Trans. Electron Dev., Vol. ED-32, p. 203, 1985.

[3] F.S. Lai et al., IEDM Tech. Dig., p. 513, 1985.

[4] T. Yamaguchi et al., IEDM Tech. Dig., p. 522, 1983

[5] M.L. Green, AT & T Bell laboratories, to be published.

[6] E.K. Broadbent et al., J. Electrochem. Sec., Vol. 133, p. 1715, 1986.

[7] E.K. Broadbent, Workshop on refractory metals and silicides for VLSI, San Juan Bautista, California, 1986.

A COMPARISON OF RETROGRADE AND CONVENTIONAL N-WELLS FOR SUB-MICRON CMOS CIRCUITS

A.G.Lewis, R.A.Martin, J.Y.Chen, T.Y.Huang and M.Koyanagi

Xerox Palo Alto Research Center
3333 Coyote Hill Road, Palo Alto, CA 94304, USA

Abstract. A general and direct comparison of retrograde and conventional n-well CMOS technologies is reported. The advantages of the retrograde structures in terms of packing density, isolation and short channel PMOS characteristics are demonstrated. The conventional wells offer slightly better circuit performance due to lower p+ to n-well junction capacitance. However, the major difference between the well types lies in their latchup susceptibility; here the retrograde wells have a significant advantage due to their lower sheet resistance and greater tolerance to very thin p-on-p+ epitaxial layers.

1. INTRODUCTION

The use of high energy (0.5-1MeV) phosphorus ion implantation to form retrograde n-wells for high performance VLSI CMOS circuit fabrication has been reported by a number of authors [1-4]. The advantages attributed to such wells include improvement in short-channel PMOS transistor characteristics [2,3], improved packing density [4] and reduced latchup susceptibility [2]. These features make retrograde wells very attractive for sub-micron CMOS circuit applications. However, in order to assess the usefulness of implanted retrograde wells fully it is also necessary to draw a direct and general comparison with conventional structures. This paper presents a broad comparison between independently optimised retrograde and conventional n-well technologies; the same mask sets have been used to fabricate devices and circuits in both processes, and thus the comparison is as direct as possible.

The lateral isolation achievable in the two technologies has been reported elsewhere [4], and is not considered in detail here. However, the combined effect of lateral isolation performance and latchup on packing density is discussed. The characteristics of active p-channel transistors formed in the two well types is also described, and the implications for circuit performance are demonstrated.

2. RETROGRADE AND CONVENTIONAL N-WELLS

Both the retrograde and conventional well technologies used in this work rely on relatively straightforward LOCOS [5] isolation. In the conventional well case, phosphorus ions are implanted at low energy (<200keV), a thermal drive in is performed and the field oxide is grown. The retrograde wells, however, are formed after the field oxidation. The phosphorus ion implantation energy (0.6-0.8MeV) is chosen such that the peak phosphorus concentration lies just below the field oxide - silicon interface, and only a short activation anneal is performed after the implant. In all respects other than the n-well formation, the two technologies are very similar.

Figure 1 shows simulated two dimensional doping contours for the region around an n-well edge, including n+ and p+ source / drain diffusions, for both technology types. In each case, the n-well depth under active area (the p+ diffusion) is just over 1μm, but the retrograde well is substantially shallower in the field region.

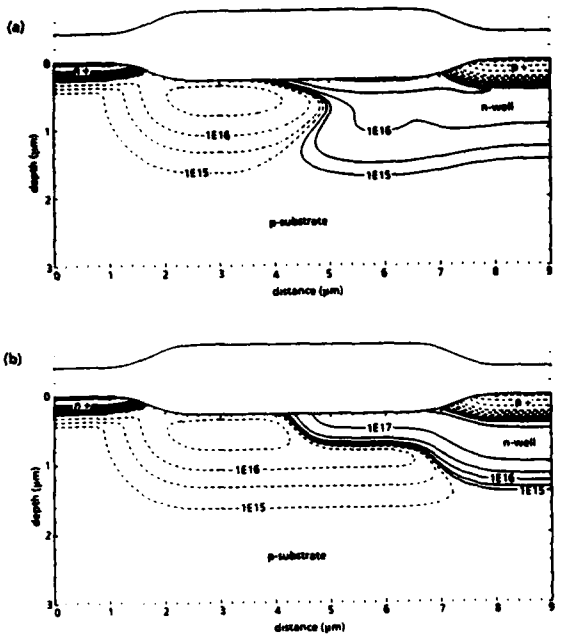


Fig.1 Simulated two dimensional doping contours at n-well edge. - - - net p-type doping —— net n-type doping. (a) Conventional n-well. (b) Retrograde n-well.

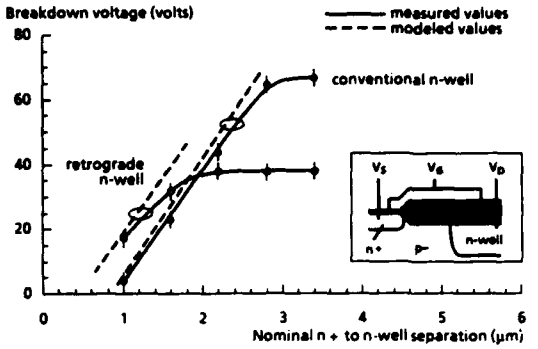


Fig.2 N + to n-well breakdown voltages. VGS = 5v ID = 10pA.

the underlying p-substrate, as illustrated in the insert, as a function of p on p + epitaxial layer thickness. On thick epitaxial layers the punchthrough voltage is sufficiently high that circuit operation is not impeded. However, if the nominal (as grown) epitaxial layer thickness is reduced to 4μm or less, the vertical breakdown voltage falls below 10 volts. Vertical isolation thus sets a minimum value on the epitaxial layer thickness, and this is important for latchup immunity.

3. EXPERIMENTAL RESULTS

3.1. Lateral isolation

Figure 2 shows n + to n-well breakdown voltages as a function of mask separation. At large spacings, avalanche breakdown at the well-substrate junction dominates, and the higher doping levels in the retrograde structure produce a lower breakdown voltage. At smaller separations, however, lateral punchthrough occurs, and here the advantage of the shallower field well depth achieved with retrograde structure is clear. The breakdown voltage does not fall below 10 volts until the nominal separation is well below 1μm. This corresponds to a minimum n + to p + separation of less than 2μm [4].

3.2. Vertical isolation

One of the major problems associated with the use of very shallow conventional n-wells is that of vertical isolation [6,7]. Figure 3 shows vertical punchthrough voltages measured between a large p + diffusion and

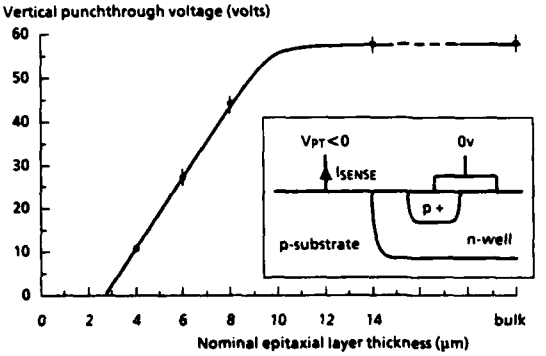


Fig.3 Vertical p + to p-substrate punchthrough voltage in a *conventional n-well*. Sense current = 2fAμm^{-2}, well depth = 1.2μm, well doping = 3x10^{16}cm^{-3}.

Vertical isolation in retrograde wells is much better, even at the same well depth, due to the higher doping levels that can be achieved at the bottom of the well without raising the surface concentration sufficiently to affect active PMOS device characteristics (figure 1). Thus, thinner epitaxial material can be used.

3.3. Latchup

Figure 4 compares latchup holding currents for retrograde and conventional n-wells. Data are shown for a range of nominal (as-grown) epitaxial layer thicknesses; the final thicknesses are about 2μm less in the conventional well case and about 1μm less with the retrograde wells, the difference being due to out-diffusion from the p+ substrate during the n-well drive-in required for the conventional technology. With nominally 4μm thick epitaxial layers, the conventional wells give higher holding currents due to this effect. However, the highest holding currents of all are obtained with retrograde wells and a 2μm as-grown epitaxial layer; conventional well devices cannot be fabricated successfully on such material. The retrograde structures also provide better latchup protection with thick epitaxial layers or bulk material due to their much lower n-well sheet resistance.

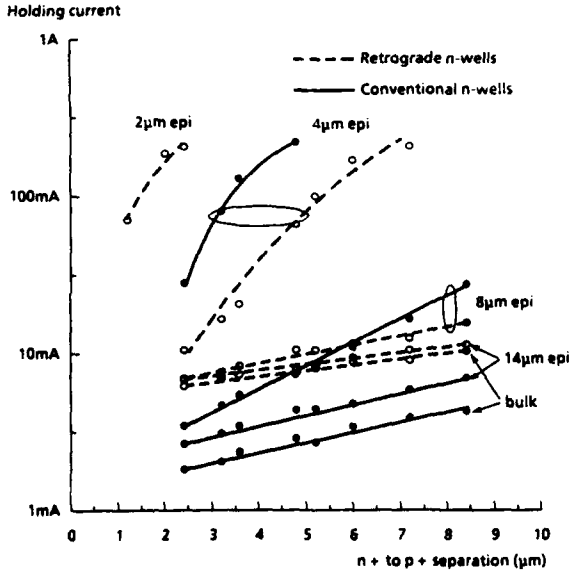


Fig.4 Latchup holding currents vs n+ to p+ separation. Diffusion width is 40μm

3.4. Packing density limitations

Figure 5 summarises the limitations imposed on packing density by latchup, lateral and vertical isolation. The boundary lines delimiting the safe operating regions are determined as follows. The minimum n+ to p+ separation is set by lateral leakage, as discussed in section 3.1. The minimum final epitaxial layer thickness is set by vertical isolation (section 3.2). Lateral and vertical breakdown voltages of at least 10 volts are required. A holding voltage of greater than 5 volts is chosen as the latchup criterion; this is determined by the final epitaxial layer thickness and is almost independent of the technology type [8]. It is clear the retrograde wells allow significantly smaller n+ to p+ separations, and that this is primarily due to their superior vertical isolation since this allows thinner epitaxial material to be used in order to maintain latchup immunity.

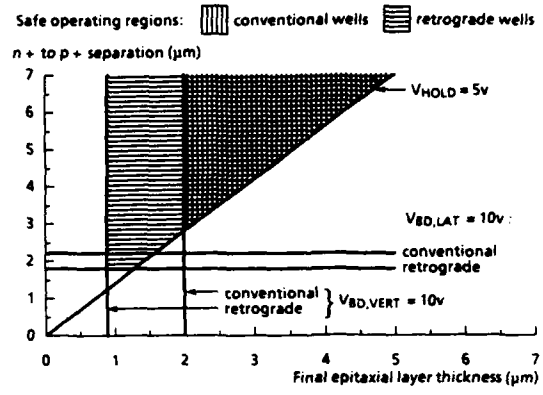


Fig.5 Safe operating regions for conventional and retrograde n-well technologies.

3.5. Active p-channel device performance

In the conventional n-well technology, the surface well doping is largely determined by the need to maintain vertical isolation, and little freedom to optimise the well profile for PMOS device performance exists. With retrograde wells, however, the deep, high energy well implant achieves vertical isolation without raising the surface doping significantly. A separate implant can thus be added and tailored for optimum p-channel performance. This greater design freedom means that good small geometry PMOS performance is more easily achieved in the retrograde wells, as illustrated in figure 6. The extra margin against PMOS punchthrough is particularly useful for sub-micron designs where hot electron induced punchthrough [9] becomes a major concern.

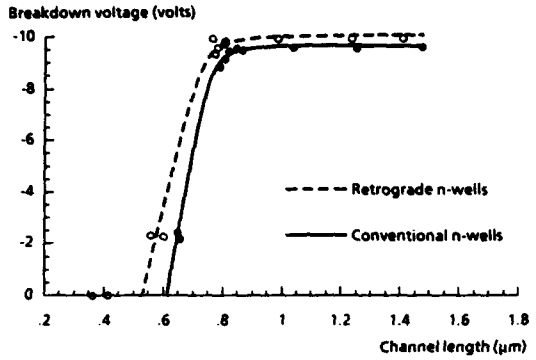


Fig.6 PMOS punchthrough voltages. $V_{GS} = V_T + 1v$ $I_D = 100pA$.

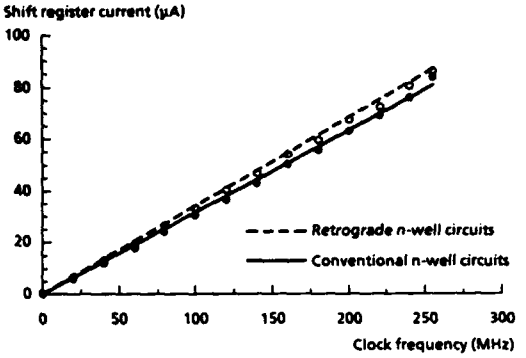


Fig.8 Ten stage static shift register supply current at $V_{DD} = 5v$

3.6. Junction capacitance

The higher doping levels in retrograde n-wells which give superior latchup immunity and vertical isolation in comparison to conventional structures also increases p+ to n-well junction capacitance. Measured zero-bias values are about 0.95fF/μm² and 0.45fF/μm² for the retrograde and conventional wells respectively.

3.7. Circuit performance

The main difference in terms of circuit operation arises because of the difference in p+ to n-well junction capacitance noted above. Figure 7 shows simple ring oscillator stage delay, while figure 8 shows the current consumption of a 10 stage static shift register as a function of operating frequency. In both cases, the conventional well circuits offer 5-10% better performance.

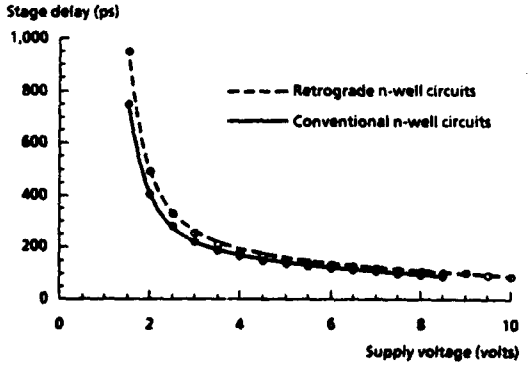


Fig.7 Ring oscillator stage delay. F.O. = 1

4. CONCLUSIONS

The comparison reported here shows that retrograde n-wells have some advantage over conventional structures in terms of packing density due to the smaller n+ to p+ separation which can be achieved with the shallower field region well depth. The implanted structures also offer greater freedom to optimise p-channel doping profiles, leading to better short-channel PMOS characteristics. On the other hand, the conventional structures offer slightly superior circuit speed and power dissipation due to lower p+ to n-well junction capacitance. However, the main difference between the technologies lies in their immunity to latchup. Here, the retrograde wells are significantly better for two reasons: firstly, the n-well sheet resistance is lower, and secondly the vertical isolation is better, allowing thinner epitaxial material to be used.

References

[1] Y. Taur et al. IEEE Trans. Electron Devices, vol ED-32, no.2, pp 203-209, 1985.
[2] R.A.Martin et al. IEDM Technical Digest, 1985, pp 403-6.
[3] F.-S.J.Lai et al. IEEE Trans. Electron Devices, vol ED-33, no.9, pp 1308-1319, 1986.
[4] A.G.Lewis et al. IEEE Trans. Electron Devices, vol ED-34, no.6, pp 1337-1345, 1987.
[5] J.A.Appels et al. Philips Research Report, vol 25, no. 2, pp 118-132, 1970.
[6] A. G. Lewis et al. IEEE Trans. Electron Devices, vol. ED-30, pp. 1680-1692, 1983.
[7] A.G.Lewis et al. IEEE Electron Device Letters, vol EDL-8 no.3, pp. 107-109, 1987.
[8] A.G.Lewis et al. 1986 Sym. on VLSI Tech. Digest, pp. 23-24.
[9] M.Koyanagi et al. IEEE Trans. Electron Devices, vol. ED-34, pp. 839-844, 1987.

IMPACT OF S/D-PREAMORPHIZATION ON CMOS PERFORMANCE

C. Mazuré, J. Winnerl, F. Neppl

Siemens AG, Corporate Research and Development, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, FRG

Substrate amorphization prior to Source/Drain implantation is used for shallow junction fabrication. The impact of preamorphization on CMOS Performance is investigated. Results for a 1.5µm double well CMOS Technology with phosphorus and boron drains are presented. The influence of preamorphization on the transistor characteristics, speed and latch-up hardness is discussed.

1. INTRODUCTION

Amorphization of Si prior to B implantation ist known to reduce the p^+n-junction depth. In addition, the electrical activation of B is improved. So far, investigations have focussed on the characteristics of the resulting p^+n-junction.

In this paper, our results of preamorphized p^+-diffusion regions are briefly summarized. For the first time, also the impact of preamorphization on phosphorus doped n^+-diffusion regions is discussed. The main part of the paper, however deals with the influence of the corresponding effects on CMOS device characteristics. The results are based on a 1.5µm double-well CMOS technology with B- and P-doped S/D-regions for p- and n-channel devices, respectively, but are also applicable for sub-µm technologies.

2. S/D-FORMATION

Amorphization of the S/D-regions was achieved by a blanket high dose Si implant prior to S/D-implantation. In order to avoid recrystallization effects due to heating during the Si-implantation, the substrate temperature was kept below 70°C. P and B implants of a dose of $5x10^{15}cm^{-2}$ were used for the formation of n^+- and p^+-diffusion regions. Recrystallization and dopant activation were done by a 40 minutes anneal at 900C°.

Fig. 1+2 characterize the effect of preamorphization on the p^+-diffusion regions. In Fig. 1 SIMS-profiles of the p^+-diffusion regions with and without preamorphization are compared. Fig. 2 shows the dependence of junction depth x_j and sheet resistance R_{p^+} on the depth x_{si} of the amorphous zone. Preamorphization has three effects on the p^+-diffusion regions:

First, the junction is considerably shallower. Reasons are the elimination of channeling and the

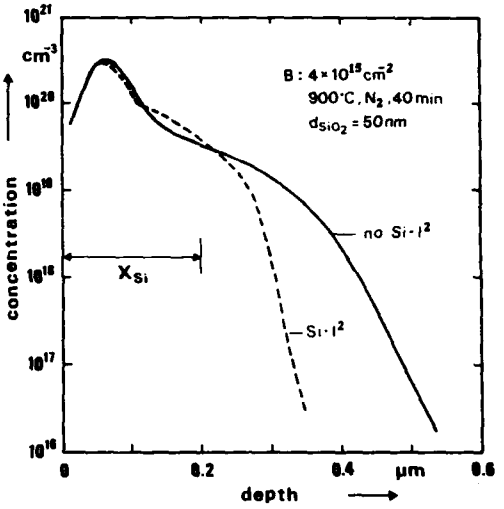


Fig. 1: SIMS profile for boron with and without preamorphization.

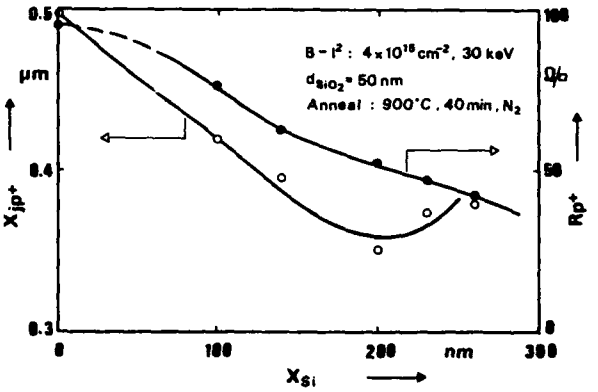


Fig. 2: Junction depth x_{jp^+} and sheet resistance R_{p^+} vs. depth x_{si} of the amorphized zone

reduction of B diffusion. As can be seen from Fig. 2, x_j depends strongly on x_{s1}, there is a minimum around x_{s1}=200nm. For significantly smaller x_{s1}-values., channelling is no longer suppressed. For higher x_{s1}-values x_j increases probably because B diffusion is enhanced by the increasing density of point defects.

Second, the activation of the implanted B is enhanced in the preamorphized samples (see Fig. 1), because there is less B clustering. This in turn reduces the sheet resistance of the p^+ S/D-regions (Fig. 2) and particularly increases the electrically active surface concentration.

Third, the amount and local distribution of the residual damage changes. In standard B-implanted p^+-regions there is a high defect concentration located at the maximum of the as-implanted B profile. With preamorphization the density of damage is considerably reduced. The residual damage is concentrated around x_{s1} (Fig. 1), which is closer to the pn-junction. Since $x_{s1}<x_j$ in all cases (Fig. 2), no impact on diode leakage current was found as expected.

The investigation of the impact of preamorphization was extended to phosphorus doped n^+-diffusion regions. n^+-diffusion regions have to be taken into account for applications in CMOS technology. Fig. 3 compares the phosphorus profiles with and without preamorphization after annealing. Similar to the case of B, the junction depth is reduced. This effect was significant for

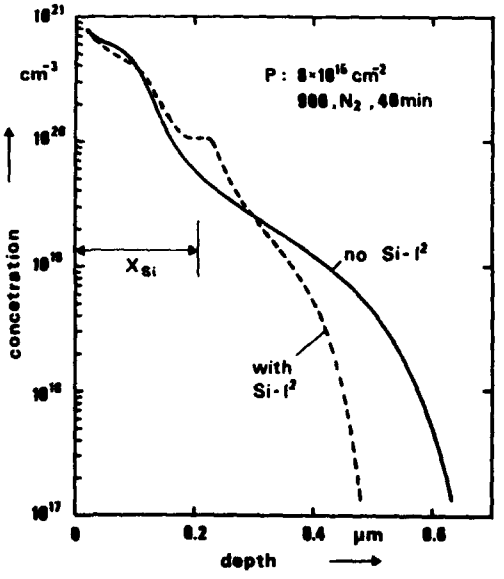


Fig. 3: SIMS profile for phosphorus with and without preamorphization.

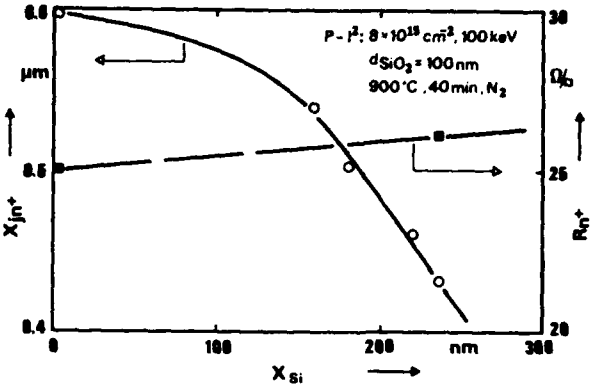


Fig. 4: Junction depth x_{jn^+} and sheet resistance R_{n^+} vs. the depth x_{s1} of the amorphized zone.

$x_{s1}\geq$200nm (Fig. 4), where the amorphized region extends beyond the region of high P concentration (Fig. 3). The accumulation of P in the residual damage zone around x_{s1} suggests, that mainly the fast vacancy related P diffusion in the tail region is retarded.

Since the high dose P implant alone also leads to amorphization of Si, no enhancement of electrical activation of P was obtained by the preamorphization. Consequently no significant effect on the n^+-sheet resistance and on the surface concentration was observed.

For the n^+-regions with and without preamorphization there is little residual damage. With preamorphization, however, the residual damage zone is closer to the junction. This is because x_j is reduced and because the damage zone is located around x_{s1} rather than around the zone of high P concentration. Again no deterioration of the junction quality was observed.

3. DEVICE CHARACTERISTICS

For device fabrication the preamorphization conditions were chosen to minimize x_{jp^+}. x_{jp^+} was reduced from 0.55µm to 0.35µm, x_{jn^+} from 0.6µm to 0.43µm.

The main advantage of the reduction of the junction depths of the S/D regions is an improvement of the short channel behaviour of the MOSFETs. Figs. 5+6 show the threshold voltages V_T of n- and p-channel transistors as a function of the effective channel length L_{eff} for S/D-formation with and without preamorphization. For both channel types the minimum channel length L_{min} with long channel behaviour is reduced by about 30% when preamorphization is used. A similar relative reduction of L_{min} is exptected for sub-micron devices, too. A major concern in

standard CMOS technologies is the insufficient subthreshold behaviour of the p-channel transistor at short channel lengths. Fig. 7 shows that due to the reduction of x_{jp^+} by preamorphization a sufficiently low subthreshold swing can be obtained for significantly lower channel lengths than without preamorphization.

A drawback of the reduction of the junction depths was the increase of substrate current I_{sub}: 20% for the n-channel and 100% for the p-channel. I_{sub} was confirmed to increase with decreasing x_j, as it is expected from simple models for the maximum electric field in the channel, which determines I_{sub}. Therefore, may be for optimum overall device characteristics a compromise has to be made between x_j reduction and I_{sub} increase.

A comparison of preamorphization with the standard technique to obtain shallow p^+n-junctions by BF_2-implantation yields the following advantages for preamorphization: no fluorine related damage, lower parasitic resistances, shallower junctions. The increase of the

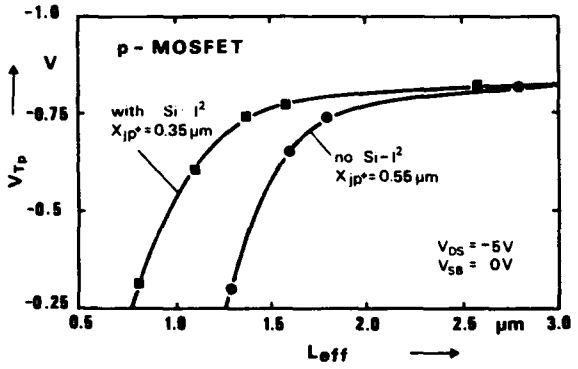


Fig. 5: p-MOS threshold voltage V_{Tp} with and without Si-I² vs. channel length.

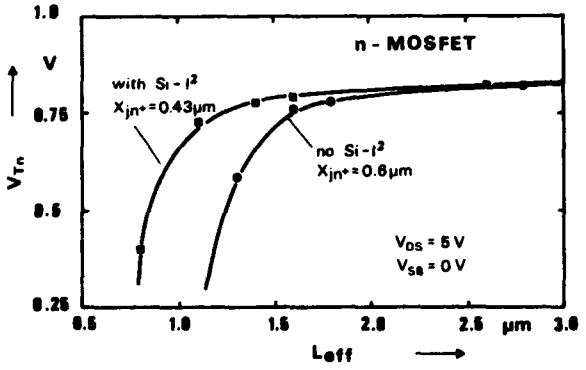


Fig. 6: n-MOS threshold voltage V_{Tn} with and without Si-I² vs. channel length.

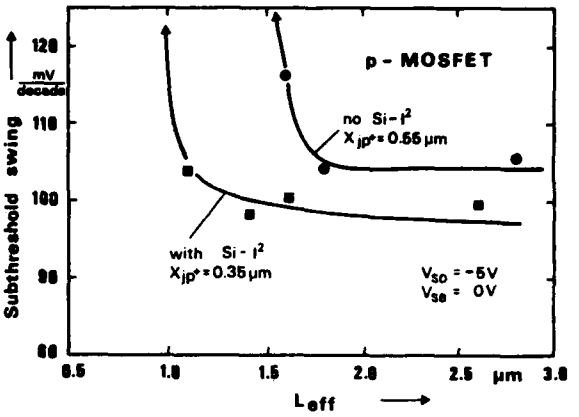


Fig. 7: p-MOS subthreshold swing with and without Si-I² vs. channel length.

substrate current with respect to only B-implanted S/D regions is no particular effect of the preamorphization but follows from the reduction of x_j and therefore is equally observed for BF_2 implanted junctions corresponding to the x_j-value obtained.

Depending on the profile of the background doping, the reduction of x_j can cause changes of the junction capacitances, which might affect circuit speed. In the present case the n-well is rather shallow and the concentration of the well doping steadily increases towards the surface. This represents the worst case, since a x_j reduction causes an increase of the junction capacitance C_j because of increasing background doping. This was confirmed by measurements of the junction capacitances of p^+n-diodes with and without preamorphization. The difference is significant at low reverse voltages only, since at high voltages the depletion region extends into the region of constant background doping underneath the shallow n-well. For the n^+-junctions no C_j change was observed. This is expected from the fact, that for the deep p-well the doping concentration is rather constant close to the surface. It was confirmed by gate delay measurements with 3-input NAND/NOR ring oscillators, that the increase of C_j at low reverse bias had no noticable effect on circuit speed. This can be understood by considering that junction capacitance is only a minor part of the total capacitive load. Other doping profiles would lead to other types of changes of the junction capacitances. It is, however, generally believed, that these changes have only minor effect on circuit speed. This holds particularly for scaled technologies, where wiring capacitance dominates.

The better electrical activation of boron reduced the p^+-sheet resistance to 60% of its value without preamorphization. The higher surface concentration leads to a factor of 2 decrease of the contact resistance between metallization and the p^+-areas. Thus the parasitic series resistance of the p-channel devices is considerably diminished. This becomes increasingly important for sub-micron technologies where parasitic series resistances begin to limit the current of the short channel transistors.

The shift of the residual damage zone towards the pn-junction which was observed for the preamorphized diffusion regions did not affect the diode leakage current, but significantly improved the latch-up hardness. This is attributed to the reduction of the current gains of the parasitic bipolar transistors the emitters of which are the S/D-regions of the MOS-devices. Fig. 8 shows for the parasitic n^+pn transistor, that the current gain steadily decreases with decreasing spacing between the damage zone and the pn-junction. For the optimum conditions chosen in our technology, the current gain is

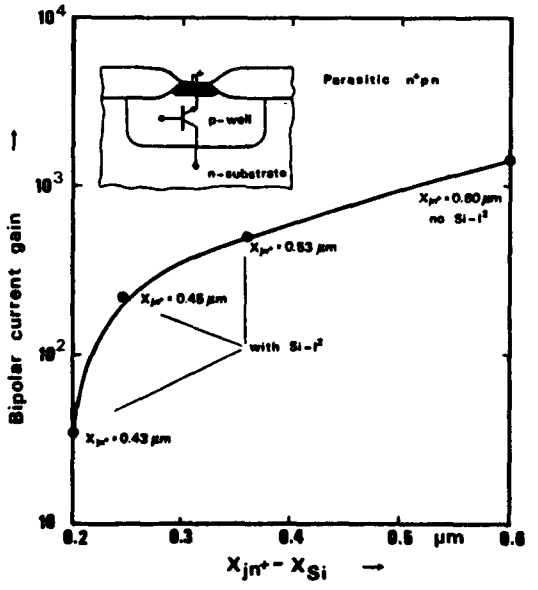


Fig. 8: Parasitic npn bipolar current gain depending on the spacing between x_{Si} and x_{jn^+}.

reduced by nearly two orders of magnitude. Similar results were obtained for the parasitic p^+np-transistor. From this it is concluded, that the small spacing between the residual damage zone in the case of preamorphization, which is a zone of high minority carrier recombination rate, and the base-emitter junction enhances the base current and thus reduces the emitter efficiency. This leads to the small current gains of the parasitic bipolar transistor in case of preamorphization, which considerably improves the latch-up hardness. Consequently preamorphization allows to reduce the critical n^+/p^+-spacing of CMOS structures without trade off with respect to latch-up hardness. For the technology presented here, the n^+/p^+-spacing could be reduced from 6µm to 5µm.

4. CONCLUSION

The impact of preamorphization on CMOS performance was evaluated for a 1.5µm technology. A considerable improvement of the short channel behavior of the MOS devices has been demonstrated. In scaled devices the phosphorus doped n^+-S/D-regions will be replaced by LDD-structures. Consequently in the future preamorphization might only play a minor role for n-channel transistors, e.g. by reducing the lightly doped phosphorus tail underneath the As. In contrast, the p-channel results will be useful also for sub-micron technologies. It has to be investigated yet, how the I_{sub}-increase affects device reliability. There might be a trade off between optimum short channel behavior and sufficient device reliability. The improvement of latch-up hardness caused by the redistribution of the residual damage of the S/D-regions will be present also in scaled devices, even if for the n-channel transistors LDD structures are used, because a similar reduction of the spacing between damage zone and junction is expected as in phosphorus doped n^+-regions.

In summary, preamorphization proved to be a useful technique to optimize future device structures.

REFERENCES

/1/ K. Yamada, M. Kashiwagi and K.Taniguchi, Jpn. J. Appl. Phys. 22-1, 157 (1982)

THE INFLUENCE OF TRENCH ISOLATION ON SUB-MICRON TRANSISTORS

M C Roberts, D J Foster, P H Bolbot and P L Medhurst

Plessey Research Caswell Limited
The Allen Clark Research Centre
Caswell, Towcester, Northants. U.K. NN12 8EQ

In order to achieve a high packing density, sub-micron CMOS process, a trench isolating technique with both n+ and p+ diffusions abutted directly to the trench wall has been developed. By placing both n+ and p+ diffusions against the trench, a packing density improvement greater than two can be achieved over a trench process with off-set diffusions.

In this work a twin well CMOS process with shallow sources and drains (0.2μm) has NMOS and PMOS transistors isolated by 1μm wide and 2μm deep trenches. The trenches are passivated with both a sidewall oxide and a nitride layer and filled with poly-Si. The sidewall nitride layer prevents the encroachment of the field oxide into the active area and hence allows the n+/p+ diffusions to abut directly to the trench wall. Fig.1 shows an SEM cross-sectional micrograph of the trench structure before gate conductor patterning.

The electrical characteristics of both NMOS and PMOS trench isolated transistors have been evaluated for a range of transistor geometries. Good working sub-micron transistors have been processed with electrical widths or lengths as low as 0.3μm (Fig.2). The process schedule gave unity aspect ratio gains of 90 and 25μA/V^2 and subthreshold slopes of 95 and 100mV/dec for 0.8μm NMOS and PMOS transistors respectively (gate oxide thickness ~ 200Å). The off-state leakage currents were less than 0.1pA/μm width for both NMOS and PMOS, and lateral punch-through voltages were over 7V for both NMOS and PMOS transistors each with an Leff=0.75μm. Transistors were fabricated on both n-type and p-type wafers.

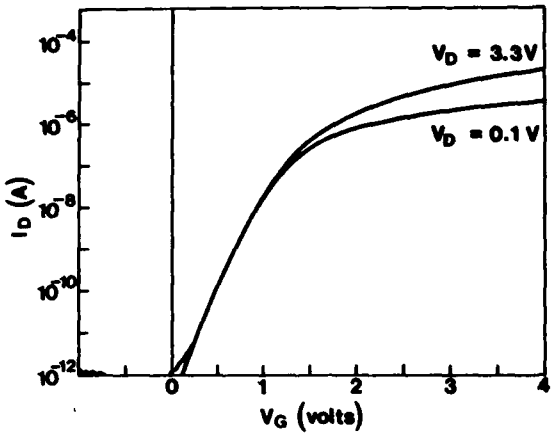


FIGURE 1

Cross-sectional SEM of the trench structure prior to poly-gate patterning.

One of the main concerns of a trench process is the ability to control the trench parasitic transistors. If the trench parasitic transistors are allowed to conduct, the quiescent power dissipation of a circuit will be extremely high.

Electrical contacts were made to the poly-Si in order that it could be used as the gate of the lateral and vertical parasitic transistors enabling the characteristics of these parasi-



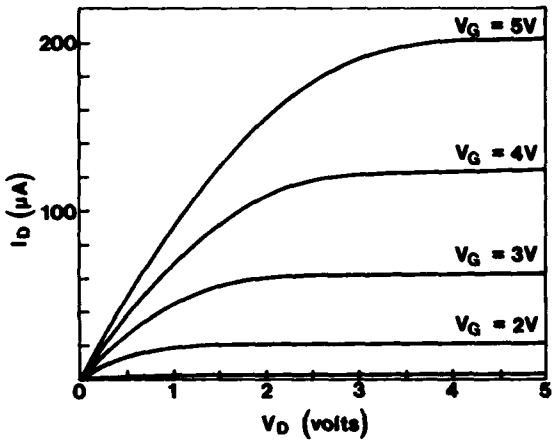


FIGURE 2

Subthreshold and I-V characteristics of ten parallel sub-micron wide transistors (W=0.3μm, L=10μm).

tics to be determined. Fig.3 shows the subthreshold characteristics of a lateral NMOS parasitic transistor. The lateral parasitic has a threshold voltage of 12V and a subthreshold slope of 875mV/dec. The trench voltage required to obtain a vertical trench leakage current of 1pA per micron length of trench and a lateral trench leakage current of 10pA per trench edge was measured. The drain potential of the parasitic transistors was kept at a potential of 5V(NMOS) and -5V(PMOS) during these measurements.

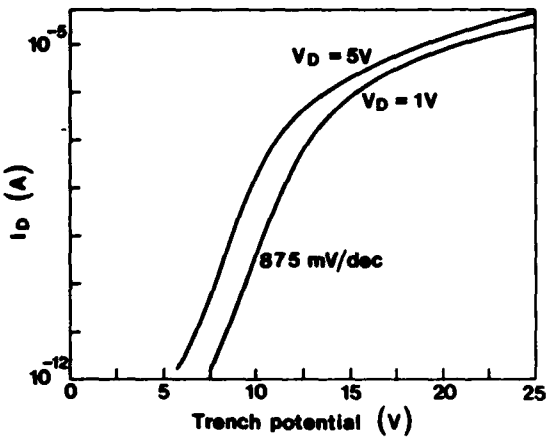


FIGURE 3

Subthreshold characteristics of a lateral NMOS parasitic transistor.

Fig.4 summarises the results for the four trench parasitics. The vertical NMOS parasitic was measured on an n-type substrate (2Ωcm) and the vertical PMOS parasitic on a p-type substrate (25Ωcm).

	Trench voltage (V)
NMOS vertical	4.5 ± 0.3 (3σ n-1)
NMOS lateral	7.5 ± 1.2 (3σ n-1)
PMOS vertical	-16.5 ± 2 (3σ n-1)
PMOS lateral	-9.5 ± 1 (3σ n-1)

FIGURE 4

Trench voltage of the four trench parasitics.

It is clear from the results summarised in Fig.4 that it is preferable to establish a process on p-type wafers since the n+ to n- substrate vertical NMOS parasitic which has the lowest turn-on voltage will be eliminated.

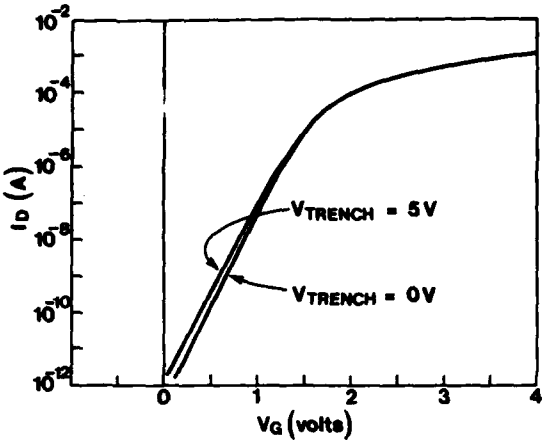
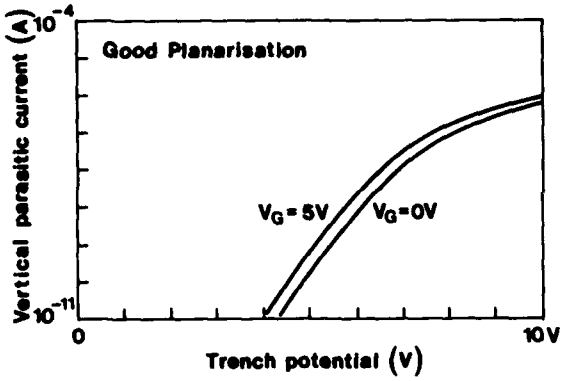


FIGURE 5

Dependence of the subthreshold characteristics of a 20/1 NMOS transistor on trench potential.



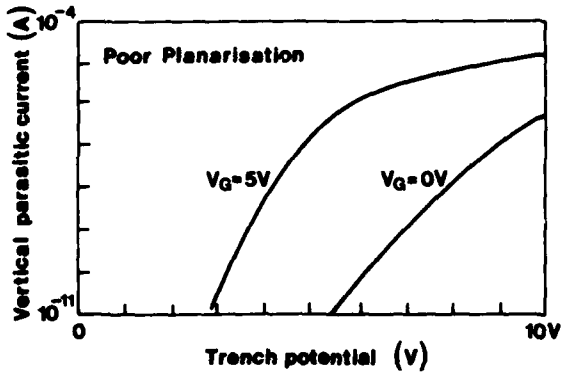


FIGURE 6

Dependence of the vertical parasitic current on the gate potential (V_{sub}=5V).

As well as characterising the parasitic transistor directly, the subthreshold characteristics of the top transistor were measured as a function of trench potential. Fig.5 shows that there is only a minimal shift in the subthreshold characteristics of a 20/1 NMOS transistor as the trench potential is increased to 5V.

Another concern of the trench process is the ability to maintain adequate planarisation. If the poly-Si gate is allowed to extend over the edge of the active area then the vertical parasitic current will be dependent on the gate potential (Fig.6). However, by maintaining adequate planarisation, this effect can be largely eliminated.

As well as characterising the trench parasitics at room temperature, full characterisation has also been carried out on all the trench parasitics in the temperature range 20°C to 125°C. Fig.7 shows the subthreshold characteristics of the NMOS lateral parasitic for a range of temperatures. There is a shift of 3.5 volts in the threshold voltage between 25°C and 125°C and a corresponding degradation in the subthreshold slope.

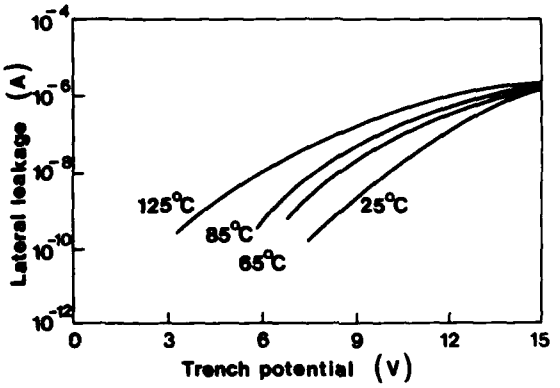


FIGURE 7

Lateral trench parasitic leakage as a function of temperature.

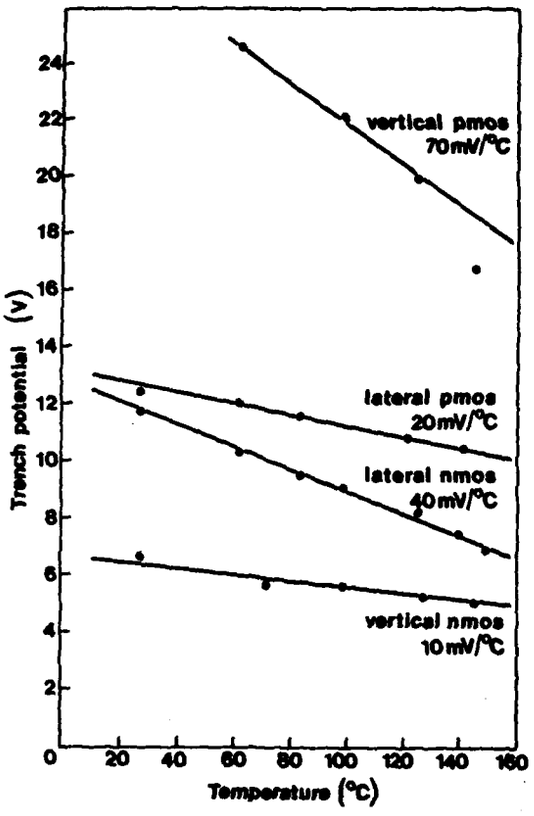


FIGURE 8

Dependence of the trench potential on temperature for the four trench parasitics (vertical current 3nA/μm, lateral current 50nA/trench edge).

Fig.8 summarises the temperature dependence of the four trench parasitics.

The results obtained show that the trench parasitics will not lead to unacceptable leakage currents even at elevated temperatures since in a real circuit formed on p/p+ epitaxy, the trench will sit at a potential which will be close to the substrate potential of 0V.

To conclude, a CMOS trench process has been developed which has both n+ and p+ diffusions abutted directly against the trench sidewall. The twin wells and the trench structure have been optimised so that the trench parasitics do not lead to problematic leakage currents. Sub-micron CMOS transistors have been fabricated with good electrical performance showing that trench isolation has potential for scaling to 0.5μm dimensions and beyond.

The authors thank the Alvey Directorate for their financial support.

Electrical evaluation of SWAMI structures

T.Cavioni and F.Gualandris
SGS Microelettronica,Central R&D, Via C.Olivetti 2
20041 Agrate Brianza (MI)-Italy

The SWAMI electrical properties have been studied focusing over the subthreshold voltage, junction leakage and latch-up phenomena. The data below reported have been compared with the equivalent LOCOS processed structures.

1 Introduction

The SWAMI (Side Wall Masked Isolation) technique has been recently proposed [1],[2],as a suitable isolation approach for the new VLSI and ULSI devices. It is interesting to remember that the SWAMI structure, beside some electrical advantages, presents less than 10% of the field oxide thickness sticking out from the silicon surface, compared with about 62% of a standard LOCOS isolation. Due to the special configuration,i.e. the recessed silicon etch, it is possible to expect a larger stress induced by the oxidation in the silicon itself . As already reported [2] it is possible to obtain the SWAMI structure by either single silicon etch or double silicon etch. The two silicon etch procedure it is recognized to be more complex from a manufacturing point of view but give better results in term of leakage current. We evaluated the feasibility of these SWAMI structures, and the high voltage TEM picture of both single silicon etch and double silicon etch are below reported. According to other authors [1], we focused our electrical evaluation on the double etch structure because recognized to be advantageous.

By comparing the TEM pictures it is realizable that the double etch approach induces a lower degree of damage in the silicon lattice. The electrical measurements performed on ours isolation structures almost confirm the reduced lattice defectivity and exhibit a good isolation properties.

2 Samples Preparation

All the samples have been processed on < 100 > oriented P-type silicon wafers with a resistivity of 1.7-2.5 ohm cm . The samples preparation process is conformable with the status of the art in the field. In the figure 1 is reported a generic process flow and the specific film thickness of ours samples. It has to be emphasized that in order to etch the single crystal silicon we choose a double frequency reactor . The $SF_6 - Cl_2$

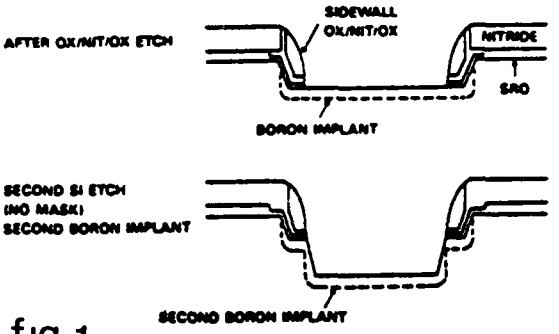


fig 1

Fig. 1) The generic process flow is here reported. The first pad oxide and nitride were 350Åand 1100Å. As reported in the text,the silicon etch was for about 500 Å. The CVD,nitride and oxide were respectively 3000 Å, 900Åand 200Å.The second silicon etch was for about 1000 Å.

plasma generated by double frequency,13.56 MHz and 100 kHz,has been found useful in order to minimize the pattern sensitivity.In other words, the double frequency configuration enhance the uniformity of the silicon etch. This is not trivial with regard to the shallow silicon layer to be etch; respectively 500 Åand 1000 Å. As mentioned above,we investigated the electrical performances of the double etch SWAMI structure. This process has been realized using also a double Boron implant respectively performed at 40 keV with a dose of $4 \cdot 10^{12} ions/cm^2$ and at 100 keV with a dose of $4 \cdot 10^{12} ions/cm^2$. The double Boron implant has been performed with the aim to prevent prevent a lateral leakage between the source and drain diffusion below the polysilicon gate. The CVD oxide as well as the second pad oxide and the second nitride have been etched in the same batch hexode reactor using a $CHF_3 - O_2$ plasma in RIE configuration. We used to this purpose a single step process in order to accomplish the theoretical requirement of a zero overetch in the wall mask generation. The field oxide, 6000 Åthick has been grown by steam technique at $920°C$ with a 3% HCl. In order to run the subthreshold measurements we patterned on the substrates two micron length N-channel parasitic transistors. The temperature influence over the junction leakage current has been studied by measuring the leakage at $25°C$ and $150°C$. A comparison between the leakage current obtained with a standard LOCOS and SWAMI is also reported. As far as LOCOS measurements are concerned we refer to the same two micron N-channel parasitic transistor processed at the status of the art. The area and perimeter diodes N+ versus P isolation and P+ versus N well have been used to check the bulk damage induced by our structure.
The area diodes were 198400 μm^2 large, the perimeter ones were 25344 μm long and 55200 μm^2 large.

3 Results and Discussion

The pictures 2 and 3 show respectively the single silicon etch and double silicon etch SWAMI structures as detected by HV-TEM (High Voltage Transmission Electron Microscopy). It is realizable that picture 3 shows a lower degree of lattice damage. We sow a lattice damage also on the double etch process but it was found at deeper level in the bulk, at about 2 μm, as show in picture 4. The bird's beak extension can be adjusted by appropriate pad oxide to nitride thickness ratio as in LOCOS technique. The bird's beak as well as the rough upper surface of the structure show in picture 3 are not representative of what we can find on the real device. This is expected to be better, because we do not performed on the TEM samples the sacrificial reoxidation process instead used on a ordinary device process.
The figure 5 shows the subthreshold characteristic as measured on our 2μm length parasitic N-channel transistor. We report on the same figure also a subthreshold characteristic of a similar parasitic transis-



Fig. 2) The HV-TEM picture of single silicon etch SWAMI .It is evident the lattice damage below the bird's beak region.



Fig. 3) The HV-TEM picture of double silicon etch SWAMI.A detail of the bird's beak region is reported.



Fig. 4) The HV-TEM picture of double silicon etch SWAMI. The view over a large region shows a sort of lattice damage in the deep bulk. It has to be emphasised that we did not perform on the TEM samples any thermal treatment.

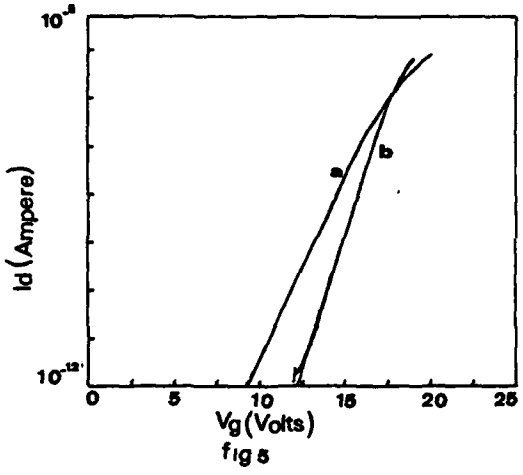


Fig. 5) The subthreshold curves for both SWAMI and LOCOS 2 μm N-channel parasitic transistors. The curve a refers to LOCOS,the curve b to the SWAMI.

tor realized using the LOCOS structure instead of the SWAMI. Based on these measures we can affirm that the defectivity induced by the SWAMI as well as the junction leakage are satisfactory and close to the values measured on LOCOS structures. In order to evaluated more accurately the induced defectivity we run some junction leakage measures using both area and perimeter diodes. In figure 6 are report the current leakage curves of N+ versus P isolation area and perimeter diodes. The upper curve refers to the area diode, the other to the perimeter one. Similar plots

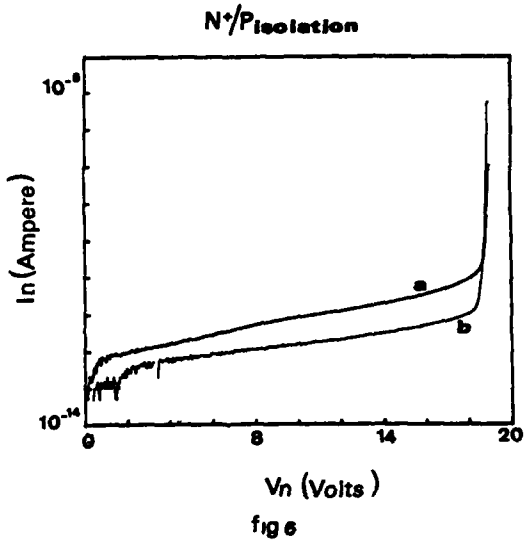


Fig. 6) Leakage current as measured on N+-P isolation area and perimeter diodes. These measures have been performed at 25°C . The curve a refers to the perimeter diode, the curve b to the area one.

but as obtained on the P+ versus N well diodes are reported in figure 7. To make a comparison the junction leakages usually measured on LOCOS structures are reported in the following table I. We report the leakage value as measured at ± 7 V and at ± 13 V .

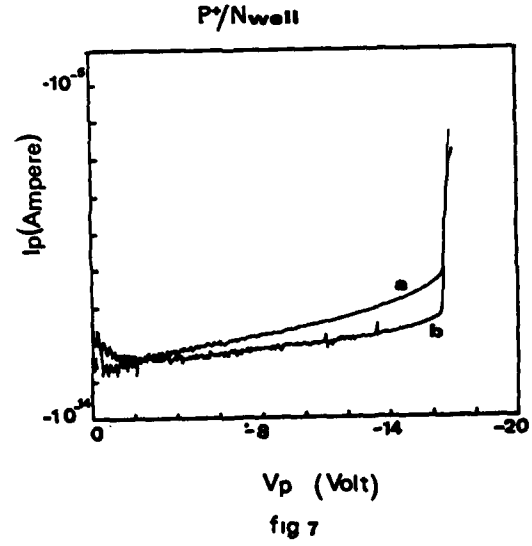


Fig. 7) Leakage current as measured on P+-N Well area and perimeter diodes. These measures have been performed at 25°C . The curve a refers to the perimeter diode, the curve b to the area one.

TABLE I

	± 7 V	± 13 V
AREA DIODES		
P+/N Well	$1.0 \cdot 10^{-12}$ A	$1.5 \cdot 10^{-12}$ A
N+/P isolation	$2.0 \cdot 10^{-13}$ A	$6.0 \cdot 10^{-13}$ A
PERIMETER DIODES		
P+/N Well	$5.0 \cdot 10^{-12}$ A	$1.0 \cdot 10^{-11}$ A
N+/P isolation	$3.0 \cdot 10^{-13}$ A	$1.3 \cdot 10^{-12}$ A

The figure 8 shows the junction leakage measured at 150°C on P+ N well and N+ P isolation area and perimeter diodes. Also these measures confirm the good quality and the low defectivity induced by SWAMI structure. The last measurements set was devoted to evaluate the consistency of the latch up phenomena on SWAMI processed samples. The figures 9 and 10 show the latch up as measured respectively on SWAMI and LOCOS structures. These measures have been performed by injecting current from the P+ emitter in N well diffused region,by keeping the N well at ground and the substrate at - 5 V. The trigger current for the latch up was found, at $6.3mA$ on SWAMI processed samples and $6.1mA$ on the LOCOS processed samples.

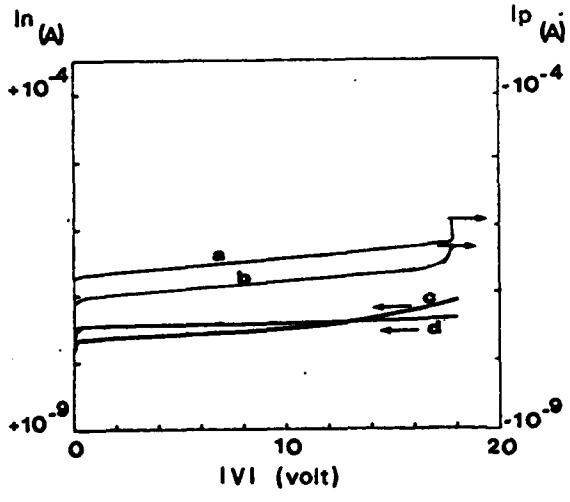


fig 8

Fig. 8) Leakage measures at 150°C . The curves a and b refer to P+-N Well diodes respectively perimeter and area. The curves c and d refer to N+-P isolation diodes respectively perimeter and area.

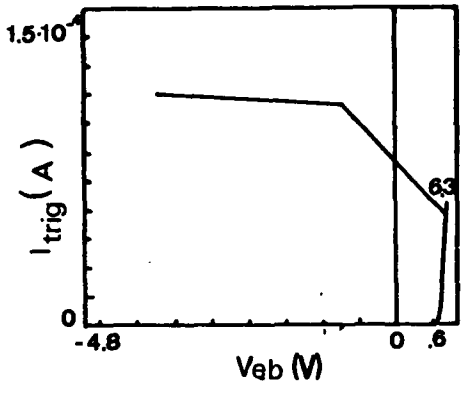


fig 9

Fig. 9) Latch-up phenomena as detected on SWAMI processed sample.

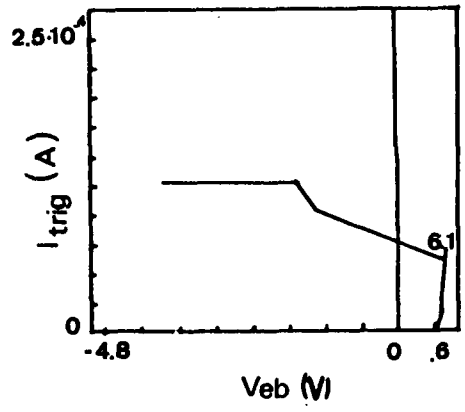


fig_10

Fig. 10) Latch-up phenomena as detected on LOCOS processed sample.

4 Conclusions

Based on ours results we can conclude that the double silicon etch SWAMI structure, is to be considered a good isolation alternative to the standard LOCOS.
As far as the electrical characteristic are concerned no significant differences with respect to the LOCOS have been found.
In spite of the recessed isolation ,the defectivity, at least its electrical evidence, do not represent at this time a dramatic constraint. The main advantages related to this structure have been found by both the higher degree of planarization achievable after the field oxide grown, and the significant reduction of the bird's beak extension. These features obtained by inducing a minimum degree of lattice damage represent the real advantage of the SWAMI technique.
In other words, the electrical characteristics comparable to those proper of the LOCOS,coupled with the geometrical advantages, i.e. high planarization level and short bird's beak extension, make SWAMI a suitable process for the new ULSI multilevel devices.

Acknowledgment

We thank C.Claeys and J. Vanhellemont of IMEC, Belgium for the TEM analysis.
This work has been supported by ESPRIT Project 554-SPECTRE

References

1) C.W. Teng, G. Pollack, W.R. Hunter
IEEE Journal of solid state circuits 20-1, 44, (1985)
2) S. Sawada, T. Higuchi,T. Mizuno, S. Shinozaki,O. Ozawa
IEEE Transactions on Electron Devices 32,11 (1985)

Session B2.4

GaAs MESFET Reliability

Chairman: F. Fantini

Tuesday, September 15, 1987

Temperature Stability of AuGeNi Ohmic Contacts to GaAs

A. Callegari, M. Murakami, J. Baker, Yih-Cheng Shih, and D. Lacey
IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

To achieve reliable, thermally stable ohmic contacts to n-GaAs, surface preparation and thickness of the AuGeNi films must be properly chosen. In this work the contact resistance as a function of the alloying temperature cycle has been studied for different AuGeNi thickness and sputter cleaning conditions. In-situ X-ray photoemission spectroscopy (XPS) analysis of the sputter-cleaned GaAs surface showed that the As_2O_3 was removed first, leaving a sputter damaged layer of GaAs containing 0.3 - 1 nm of Ga_2O_3. If a thin As_2O_3 layer was left on the surface, the contact resistance was large and non-uniform. At the optimum sputter cleaning conditions, when 5 nm of Ni was deposited first followed by 100 nm of AuGe, 30 nm of Ni and 50/100 nm of Au, the contact resistance was low and uniform with $R_c \simeq 0.1\ \Omega - mm$. Transmission electron microscope (TEM) analysis showed that a high density of uniform NiAs(Ge) grains at the GaAs interface is responsible for the much improved uniformity and thermal stability. Spread in contact resistance is due to the β - AuGa phase contacting the GaAs.

1. Introduction

Despite a trend towards developing refractory ohmic contacts which can withstand high temperature anneals, the alloyed AuGeNi contact is still the most reliable and commonly used metallization in GaAs technology. The problems associated with this system are the poor uniformity and thermal stability of the contacts. However, process improvements [1] have led to an ohmic/GaAs interface which does not degrade appreciably after thermal treatment at 400 °C for several hours. Both surface preparation and film deposition processes are of primary importance to achieve low contact resistance with good thermal stability.

In this work, the effect on contact resistance of different surface preparations and film thicknesses will be described. It will be shown that a rather narrow process window in terms of surface preparation and film thicknesses lead to a low and uniform contact resistance with good thermal stability. An in-situ sputter cleaning of the GaAs surface and the deposition of a first layer of Ni gave best results. The strong correlation between these contacts and the interfacial microstructure will be discussed.

2. Experimental

To check the feasibility of a AuGeNi process, transmission line test (TLM) test structures are formed on the GaAs wafers. Typical process steps are outlined below [1]. Conducting channels are formed by implanting, through a photoresist stencil, SiF^+ ions at a dose of $1.5 \times 10^{13}/cm^2$ at 150 KeV. After resist stripping, the wafers are capless annealed in an arsine atmosphere at 800 °C for several minutes to activate the implant. An ohmic contact lift-off stencil is then defined using photolithography. At this stage of the process, the GaAs wafer is cleaved into chips. The chips are mounted on a RF cathode and loaded into the evaporator. An in-situ sputter cleaning process is carried out first in an O_2 discharge at low power and high pressure for a short time to remove hydrocarbons contaminants and normalize the surface to a fixed oxide layer. A mild sputter cleaning in Ar is then carried out to partially remove the oxide layer. The AuGe, Ni, and Au metal films are thermally evaporated. After lift-off, the contacts are alloyed in a furnace with continuous flow of Ar/H_2. At this stage the TLMs are ready for automated test.

3. Surface preparation

Analysis of the GaAs surface during Ar sputtering has been done [2]. In-situ ellipsometry to determine oxide thickness and X-ray photoelectron spectroscopy to determine the sputtered species reveal that first As_2O_3 is removed, leaving a sputter damaged layer of GaAs containing 0.3 - 1 nm of Ga_2O_3. At short sputtering times, the contact resistance as a function of the alloying temperature cycle was measured [3]. Contact resistances showed poor uniformity and thermal stability with $R_c \simeq 0.65\ \Omega - mm$. These poor results appear to be due to a thin layer of As_2O_3 still left on the surface. At longer sputtering times, when a thin layer of Ga_2O_3 was left on the surface contact resistances were low and uniform (Figure 1a).

4. Film deposition

Since the surface can be controlled very well by sputter cleaning [1,3], different processes can be tried where the film

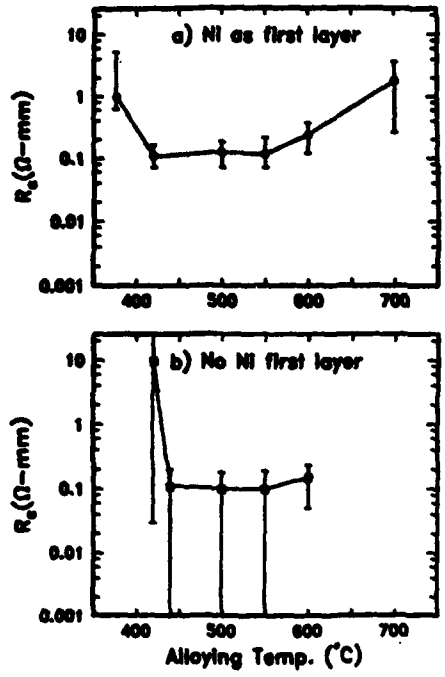


Figure 1

Contact resistance vs alloying temperature. a) Ni as first layer, b) no Ni as first layer.

thicknesses are changed. A first layer of Ni of about 5 nm [1] before the AuGe deposition, that at the beginning was thought to improve adhesion, was found to be a crucial step to obtain uniform and low contact resistance [2]. Shih et al. [4] found that it changes the sequence of the alloying reaction between AuGeNi and GaAs, which results in a large difference in the interface morphologies between specimens with and without the Ni first layer. These interface morphologies are strongly correlated with the electrical properties of the contacts. The deposition of Ni as the first layer leads to an alloyed contact formed by a two layer structure. The first layer contacting the GaAs consists of a high density of the NiAs(Ge) grains and the second layer is a homogeneous layer with large grains of the β – AuGa phase. A TEM picture of this structure is shown in Figure 2.

The contact resistance as a function of alloying temperature for this sample is shown in Figure 1a. The contact resistance is low and uniform in the temperature range of 420-550 °C. For example, at an alloying temperature of 420 °C, this sample shows $R_c \approx 0.11\ \Omega$ – mm with a total spread $\delta \approx 0.09\ \Omega$ – mm. Channel resistivities are on the order of 220 Ω/□. In these experiments about 20 TLMs are tested.

The $Ni_2AsGe/GaAs$ interface was previously suggested to be essential for low contact resistance by Kuan et al. [5]. Our present results also support this conclusion: the NiAs(Ge)/GaAs interface is important to lower the contact resistance. However, the Ge:As ratio in the NiAs(Ge) grains is not one to one. These compounds satisfy the condition that the Ge atoms diffuse into the Ga vacancy sites forming a heavily doped n^+ layer at the metal/GaAs interface. The Fermi level in the GaAs moves into the conduction band and electron transport across the interface is mainly by tunnelling through a thinner barrier height.

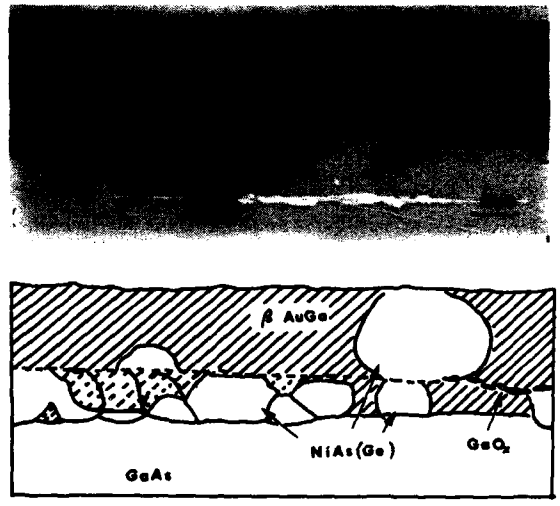


Figure 2

Bright-film image of cross sectional view of a sample with Ni as first layer after annealing at 440 °C for 2 m. The diagram below indicates the phases observed in the image area.

A much different interface morphology is obtained when no Ni is deposited first. In this case (Figure 1b), the NiAs(Ge) grains formed at the interface protrude into the GaAs substrate, rather than forming a uniform two-layers structure. Contact resistance is low but it has a large spread due to the fact that large areas of the GaAs interface are contacting the β – AuGa phase, which is a rectifying contact (Figure 3).

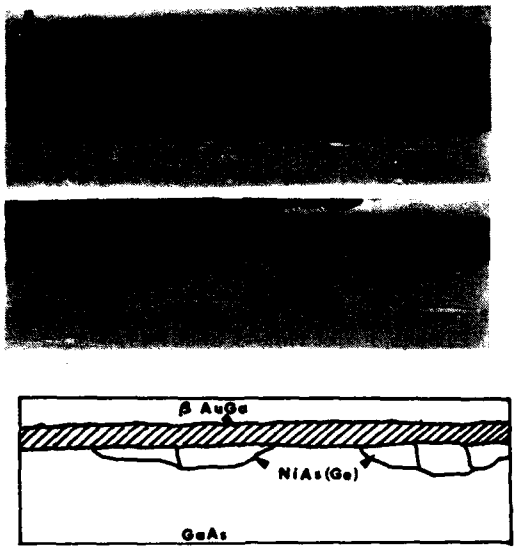


Figure 3

Bright-field image of a sample with no Ni as first layer annealed at 440 °C for 2 m at two locations, a) near a contact edge and b) at middle of a contact. In the diagram below, the images indicates the phases observed in the image area.

This type of ohmic contact has been modelled by N. Braslau [6]. Conduction is through a parallel array of germanium-rich protrusions of negligible contact resistance. Contact resistance is dominated by the spreading resistance in the semiconductor in series with them. For doping levels found in active devices, the spreading resistance which depends on semiconductor resistivity and thus inversely on doping dominates, consistent with the observed doping dependence. This widely accepted model for the AuGeNi contacts, however, cannot be strictly applied to the contacts described above, where the spacing between the protrusions has been considerably reduced and the NiAs(Ge)/GaAs interface improved.

5. Thermal stability

Figure 1a shows that after alloying at 700 °C, well above the eutectic temperature of the AuGe, the contacts were still ohmic but of poor quality. This result indicates that the contacts might show good thermal stability at 400 °C or higher temperature. When the contacts were annealed at 410 °C for 57 hours, a contact resistance $R_c \leq 0.6\ \Omega$ – mm was obtained [1].

It was found that that the deterioration of the contacts is not an increasing monotonic function with time [1]. To correlate the electrical properties to the film microstructure, the samples annealed at 410 °C after contact formation at 440 °C, were investigated using TEM. In Figure 4 segregation of the NiAs(Ge) grains is observed after annealing at 400 °C for 10 h, which reduces the contact areas between the NiAs(Ge) grains and GaAs. A very non - uniform layer was observed after 90 hours annealing (Figure 5). This deterioration is likely due to



Figure 4

TEM micrograph of the interface structure after annealing at 400 ° C for 10 hours.



Figure 5

TEM micrograph of the interface structure after annealing at 400 °C for 90 hours.

the existence of the β – AuGa phases, with a low melting point around 375 °C, which degrades contact resistance. Higher contact resistances with poor thermal stability were found when Ni or Au rich films were deposited onto the sputter cleaned GaAs surfaces.

6. Conclusions

The AuGeNi alloyed ohmic contact to GaAs is still the most widely used, due to its process simplicity. Its thermal stability, related to interface properties can be improved by sputter cleaning the surface and depositing a first layer of Ni. At 400 °C the alloyed contact is metastable, but it maintains $R_c \leq 0.6\ \Omega - mm$ after annealing for several hours. Deterioration of the contact resistance with annealing time is due to segregation of the NiAs(Ge) grains at the interface and to the existance of the β - AuGa phases with a low melting point around 375 °C.

Acknowledgments

Technical assistance from E. L. Wilkie and My Su is greatly appreciated. The authors would like to thank N. Braslau for fruitful discussion and J. H. Magerlein and J. Greiner for critical comments on the manuscript.

REFERENCES

[1] A. Callegari, E. T-S. Pan, and M. Murakami, Appl. Phys. Lett. 46, 1141, 1985.

[2] M. Murakami, K.D. Childs, J. Baker, and A. Callegari, J. Vac. Sci. Technol. B4, 903, 1986.

[3] A. Callegari, D. Lacey and E. T-S. Pan, Solid State Electron., Vol. 29, 523, 1986.

[4] Yih-Cheng Shih, M. Murakami, E.L. Wilkie, and A. Callegari, in print.

[5] T. S. Kuan, P. E. Batson, T. N. Jackson, H. Rupprecht, and E. L. Wilkie, J. Appl. Phys. 45, 6592, 1983.

[6] N. Braslau, J. Vac. Sci. Technol. 19, 803, 1981.

THERMAL STABILITY OF NON-ALLOYED OHMIC CONTACT TO n-GaAs

A. PACCAGNELLA

Dipartimento di Ingegneria, Università di Trento, 38100 Trento - Italy

A. MIGLIORI

Dipartimento di Fisica, Università di Bologna, 40126 Bologna - Italy

M. VANZI

Telettra S.p.A., Quality and Reliability Dept., Via Capo di Lucca 31, 40126 Bologna - Italy

B. ZHANG and S.S. LAU

Department of Electronic Engineering and Computer Sciences, University of California at San Diego, La Jolla, CA 92093, USA

The thermal stability of the non-alloyed GaAs/Pd/Ge ohmic contact to n-GaAs, based on the solid phase epitaxy of Ge on GaAs through PdGe, was investigated at 300°C. Annealings up to 200 hours induced an increase of the contact resistivity from the starting average value of 0.16 Ohm-mm up to 0.30 Ohm-mm. Even though no structural modification was detected by MeV Rutherford Backscattering Spectrometry measurements and Scanning Electron Microscopy observations, cross sectional Transmission Electron Microscopy analyses detected some changes at the PdGe/Ge interface in the long term annealed samples which could induce the formation of a resistive layer between Ge and PdGe.

1. INTRODUCTION

Ohmic contacts to n-GaAs are usually achieved by alloying a AuGeNi multilayer in a reducing atmosphere at high temperatures for short time (e.g., 450°C 1 min) [1]. The complex alloying process involves the formation of a liquid phase between Au and Ge, with an eutectic temperature of 356°C. A non-planar metal/GaAs interface morphology can result, as well as a poor edge definition. These morphological problems, which can be important in shallow junction devices or in high definition planar structures, can be bypassed by using a non-alloying process ensuring a planar interface, such as the solid phase epitaxy of Ge through PdGe on GaAs, as shown by Lau and coworkers [2]. A usual lift-off process is needed of the achievement of such a metallization, which gives contact resistivity values as low as 10^{-6} Ohm-cm^2 on n-GaAs, n$\sim 10^{18}$ cm^{-3}.

The purpose of this work is to investigate the thermal stability of these PdGe/Ge/GaAs ohmic contacts, in order to test their possible application as source and drain contacts in power MESFET's. Actually in such devices the temperature increase during operating life raises severe reliability problems concerning the stability of the metallization electrical properties; their degradation gives rise to a decrease in the performances of the device and eventually to catastrophic failure [3].

2. EXPERIMENTAL

2.1. Sample preparation

Semiinsulating <100> GaAs substrates with a n-doped (n$\sim 10^{17}$ cm^{-3}) active layer 0.3 μm thick were chemically cleaned with TCE, acetone and isopropyl alcohol, rinsed in HCl:H_2O 1:1 solution and eventually in deionized water, prior loading in an oil-free vacuum system. A base pressure of 2×10^{-8} Torr was reached before evaporating 50 nm Pd and 130 nm Ge in the GaAs/Pd/Ge configuration. Two different evaporation runs were performed in order to test

the repeatibility of the experimental results on different sets of samples. All samples were annealed at 325°C for 30 min in a flowing forming gas furnace (15% H_2, 85% N_2) to achieve the Ge solid phase epitaxy as shown in Section 3.1. Accelerated aging of the samples was obtained by subsequent annealings at 300°C up to 200 hours in the same furnace.

2.2. Electrical and microstructural measurements

After the metal deposition, two different sample structures were prepared by using a conventional lift-off technique in order to perform Transmission Line Model (TLM) [4] measurements of the contact resistivity ρ_t and of the average channel resistance; in the two test patterns the pad separation ranged from 2 to 42 μm and from 5 to 25 μm, respectively, and the pad width was 200 and 50 μm, respectively. The pad separations and widths were measured for each sample by using a calibrated optical microscope.

Measurement of ρ_t on patterned samples and 3 MeV $^4He^{++}$ Rutherford Backscattering Spectrometry (RBS) analyses of samples with extended metallization were performed after each thermal treatment. Cross-sectional Transmission Electron Microscopy (TEM) and surface Scanning Electron Microscopy (SEM) observations were performed only on selected samples.

3. RESULTS AND DISCUSSION

3.1. Contact formation

The interfacial structure of the metal/GaAs contact has been investigated by RBS and TEM. The RBS spectra of an as deposited and a 325°C 30 min annealed GaAs/Pd/Ge sample are shown in Fig. 1. In the annealed sample, Pd reacted with Ge resulting finally in the stoichiometric PdGe layer at the sample surface, as indicated by the relative heights of the Pd and Ge signal in the RBS spectrum. In the TEM micrograph of the annealed sample reported in Fig. 2a, the 160 nm thick PdGe layer shows a columnar structure, with grains grown perpendicularly to the surface. At the bottom of the same figure, the

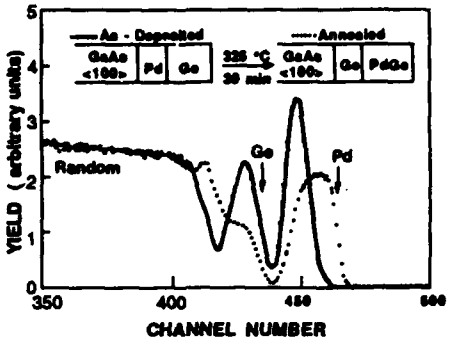


Fig. 1 Rutherford Backscattering spectra of 3 MeV ^4He ions at a 170° scattering angle for a GaAs/Pd (50 μm)/Ge (130 nm) sample before and after annealing at 325°C for 30 min.

GaAs substrate appears dotted, due to artifacts induced by the ion milling thinning of the sample for the TEM analysis. A 20 nm thick layer is present between PdGe and GaAs; by comparison to the RBS spectrum reported in Fig.1, this layer has to be attributed to Ge, epitaxially grown on GaAs as reported by Marshall et al. [2]. The Ge layer is not uniform in thickness, but it shows no extended defect such as stacking faults, previously observed in similar structures by Sawada et al. [5]. The Ge/GaAs heterointerface is planar and abrupt to within about 50 Å, while several well-shaped micrograins appear at the PdGe/Ge interface, which cannot be attributed for sure to Ge nor to PdGe. SEM observations of the contact surface after the 325°C 30 min annealing revealed a smooth and featureless morphology, with an excellent edge definition.

Therefore, the 325°C 30 min annealing induces the whole consumption of free Pd in PdGe. The residual Ge, not involved in the compound formation, is transported through Ge and grows epitaxially on the GaAs substrate, as depicted on the top of Fig. 1, due to the lower free energy of the crystalline Ge in comparison with the amorphous.

The contact resistivity of the 325°C 30 min annealed patterned samples, measured by the TLM method, gave an average value ρ_t=0.16 Ohm-mm with a data spread of 0.04 Ohm-mm over

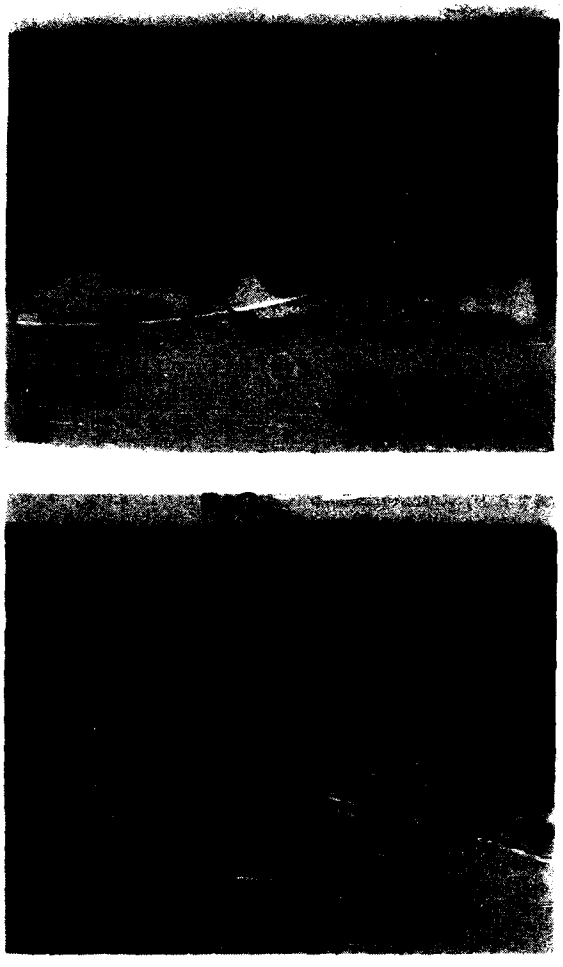


Fig. 2 Transmission electron micrograph of a cross-sectional view of a GaAs/Pd/Ge sample
a) annealed at 325°C for 30 min;
b) annealed at 325°C for 30 min and subsequently at 300°C for 188 hours.

25 different samples. Such a low contact resistivity is probably induced by the formation of a thin n^+-GaAs layer at the GaAs/Ge interface, due to Ge diffusion in GaAs, as suggested by Marshall et al. [6]. Thus a dominant tunneling conduction mechanism results at the GaAs/Ge interface, with no spiking effect in the current transport ensured by the homogeneity of such interface over the contact area.

3.2. Contact stability

The thermal stability of the GaAs/Ge/PdGe ohmic contact was investigated by subsequent thermal treatments at 300°C on different sets of samples. These annealings induced a fast increase of the contact resistivity during the first 50 hours, followed by a slow increase up to a value of about 0.30 Ohm-mm after 200 hours, as shown in Fig. 3. No increase of the

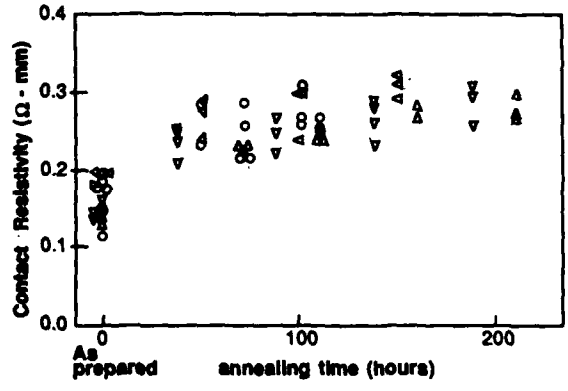


Fig. 3 Contact resistivity as a function of annealing time at 300°C, for samples previously annealed at 325°C for 30 min. Different symbols correspond to different sets of samples.

data spread has been observed after the long term treatments, even though different patterned structures, prepared in different experimental runs, were used. At the same time, no systematic increase of the channel resistance was observed in the 300°C annealed samples. RBS measurements and SEM observations on the long term annealed samples detected no noticeable variations after the prolonged thermal treatments in comparison with the as prepared (325°C 30 min) samples. No major reaction developed during the 300°C treatments involving either GaAs, Ge or PdGe.

TEM micrographs of an as prepared sample and of a sample heated at 300°C for 188 hours are reported in Fig. 2 , showing once more that no large modification affected the PdGe layer nor the GaAs/Ge interface. However, an appreciable modification of the PdGe/Ge interface appears involving the number, shape, and probably also atomic density of the micrograins,with respect to the as prepared sample. The thickness of the Ge is moreover slightly

reduced as shown in Fig. 2b.

The correlation between the PdGe/Ge interface modifications and the slight Ge thickness reduction, and the increase of the contact resistivity, is not straighforward. However, the formation of a resistive layer between PdGe and Ge, and/or a local modification of the band structure due to the narrowing of the Ge layer may be the likely origins of the ρ_t increase. Work is in progress in order to determine the crystalline structures and compositions of the micrograins at the PdGe interface.

4. CONCLUSIONS

GaAs/Pd/Ge metallization structure gives low resistivity non-alloyed ohmic contacts, based on the epitaxial growth of a Ge layer between GaAs and PdGe induced by a thermal treatment at 325°C for 30 min. A planar metal/GaAs interface is preserved after the solid phase epitaxy process, as well as a good edge definition, which make this contact useful for shallow junction and small size device applications. A very low data spread was also observed in different experimental runs, confirming the easy reproducibility and uniformity of such a low resistivity contact across a wafer. Upon 300°C annealings, the contact resistivity increased from a starting value of 0.16 Ohm-mm up to 0.30 Ohm-mm after 200 hours with no increase in the data spread. The resistivity increase can be correlated to microstructural modifications at the PdGe/Ge interface.

AKNOWLEDGMENTS

We are grateful to dr G. Donzelli and Mr P. Angione for the preparation of some of the Test Patterns used for the electrical measurements, and to the CNR-Lamel Institute of Bologna for the technical support for TEM analyses and cross-sectional specimen preparation. One of us (A.P.) is grateful to CNR-P.F. MADESS for financial support during his stay in UCSD

REFERENCES

[1] C.J. Palmstrøm and D.V. Morgan, in: Gallium Arsenide, eds M.J. Howard and D.V. Morgan (J. Wiley, N.Y., 1985) ch. 6.

[2] E.D. Marshall, W.X. Chen, C.S. Wu, S.S. Lau and T.F. Knech, Appl. Phys. Lett. 47 (1985) 298.

[3] C. Canali, F. Castaldo and E. Zanoni, Microelectron. Reliab. 24 (1984) 947.

[4] H.H. Berger, Solid State Electron. 15 (1972) 145.

[5] T. Sawada, W.X. Chen, E.D. Marshall, K.L. Kavanagh, T.F. Knech, C.S. Pain and S.S. Lau, Mat. Res. Soc. Symp. Proc. 54 (1986) 409.

[6] E.D. Marshall, B. Zhang, L.C. Wang, P.F. Jiao, W.X. Chen, T. Sawada, S.S. Lau, K.L. Kavanagh and T.F. Knech, submitted to J. Appl. Phys.

GaAs-DEVICE LIFE-TIME IMPROVEMENTS BY NEW RESULTS ON METAL-ELECTRODE FABRICATION

R.P. Gupta+, H.L. Hartnagel, K.-H. Kretschmer, R. Schütz, J. Würfl

Institut für Hochfrequenztechnik, Technische Hochschule Darmstadt
Merckstraße 25, 6100 Darmstadt, West Germany

+ on leave from Central Electronics Engineering Research Institute
Pilani (Rajasthan) 333031, India

The development of an Au-WSi_2-Ge ohmic contact on GaAs for highly stable devices is reported. Coplanar interelectrode material migration is minimized by a technology leading to oxygen-poor GaAs surfaces.

The life-times of GaAs devices, such as FETs and led's, depend strongly on the stability of their metal electrodes, i.e. Schottky-, ohmic and heat sinking contacts. Two important effects have been considered in detail here, namely

i. the long-term interdiffusive effects of metal sandwiches and

ii. the electric-field stability of electrode edges concerning in particular metal migration between closely spaced neighbouring electrodes.

Regarding the first case, it is well known that certain metals such as Au or Ni must not interact with GaAs since an exchange process involving particularly Ga causes a slow deterioration of the contact quality. Therefore, a diffusion barrier based on WSi_2 was selected by us after careful evaluation of the available experimental data and various contact sandwiches involving this barrier were fabricated. In particular, a highly stable ohmic contact was made using

(i) e-beam evaporated 100 nm Ge (as dopant after annealing at above 500° C);

(ii) 100 nm WSi_2 sputter deposited and

(iii) a layer of Au (Fig 1).

The ohmic behaviour is demonstrated by Fig. 2. Using transmission line and sidewall resistor methods, (giving both specific resistances better than $5 \cdot 10^{-5}\ \Omega cm^2$ for nGaAs with $n = 10^{16}\ cm^{-3}$, Si doped), and based on XPS studies involving Ar^+ ion etching (Fig. 3 + 4), we established that up to values higher than 460°C no modification of the contact resistance or the compositional profile could be observed, even after prolonged thermal stressing for more than 200 hours. The W profile of the XPS data remains stable with respect to the WSi_2-Ge interface up to temperatures of 510° C.

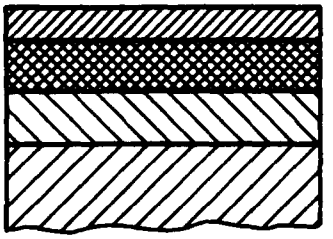


FIGURE 1
Structure of the metallization used

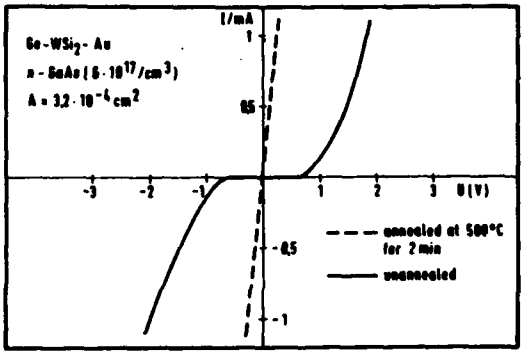


FIGURE 2
I/V characteristic of two circular patterns

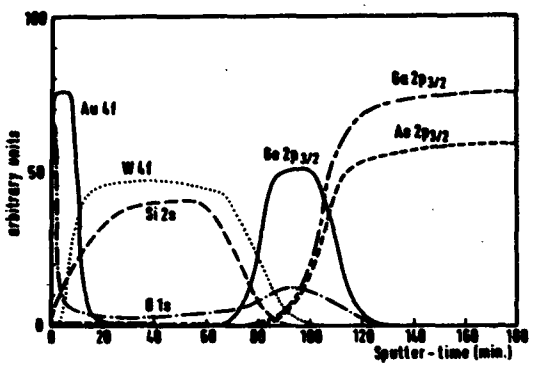


Figure 3

XPS-sputter profile of a sample annealed at 460° C for 1 hour

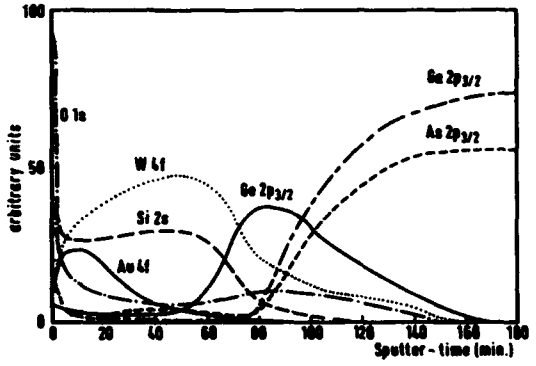


Figure 4

XPS-sputter profile of a sample annealed at 610° C for 1 hour

This work shows that it is possible to produce reliable contacts on GaAs operating even at elevated temperatures in contrast to commonly employed ohmic contacts of AuGeNi and others.

Regarding the effect (ii), interelectrode material transport in unpassivated and in PECVD-Si_3N_4-passivated GaAs planar structures was studied. Metal migration between the electrodes due to the application of an electric field was found to depend strongly in both cases on the GaAs surface condition prior to structure deposition. Alkaline etchants are found to be superior to acidic ones to terminate the treatment of the GaAs surface before the various thin-film depositions.

XPS studies have shown that the onset of material migration is directly proportional to the amount of As_2O_3 on the GaAs surface (Fig.5). The best results were achieved with an ammonia treatment of the GaAs surface, whose compositional development is given by Fig. 6. Fig. 7 shows a scanning-electron micrograph of an interelectrode short circuiting bridge. Fig. 8 illustrates the areas of interest. This MeSFET-type structure consists of two metal contacts on GaAs. The positively biased ohmic contact at the left side of the Figures is a AuGeNi-alloy, the Schottky contact at the right side consists of Al.

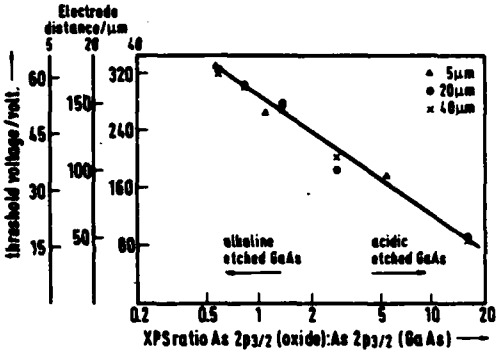


Figure 5

Threshold voltage for material migration versus XPS-ratio As (oxide) : As (GaAs)

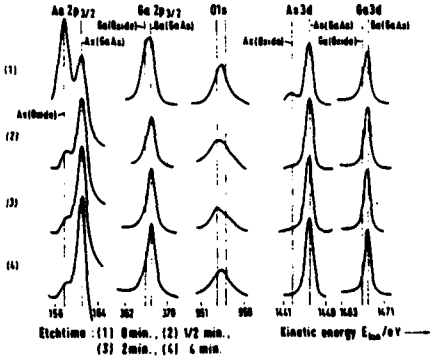


Figure 6

XPS-spectra of (100)-GaAs surfaces treated by NH_4OH (25 %) : H_2O_2 (30 %) : H_2O with the ratio 2 : 1 : 300

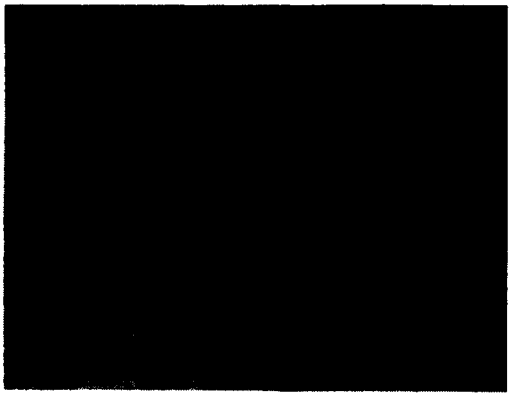


Figure 7

SEM micrograph of interelectrode short-circuiting bridge

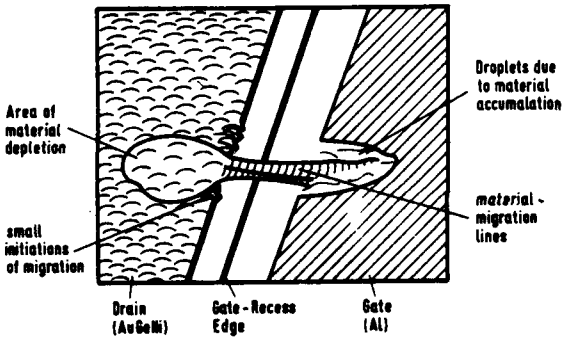


Figure 8

Illustration of interelectrode short-circuiting bridge according Fig. 7

A passivation layer does not prevent mirgration entirely. Fig. 9 shows an electrode gap consisting of two AuGeNi contacts passivated by Si_3N_4. Migration took place from the positively to the negatively biased electrode. It seems that the locality of migration is the interface between the GaAs surface and the Si_3N_4-layer. This point seems to be confirmed by other observations.

We obtained satisfactory Si_3N_4-GaAs interfaces if we applied an ammonia plasma-treatment before Si_3N_4 deposition to the GaAs surface. This treatment was performed in the same reactor and with the same conditions as for Si_3N_4 plasma deposition, but the silane input simply switched off.

It is thus possible to select an optimized technology to reduce material instabilities due to electric fields between closely spaced electrodes.



Fig. 9

SEM micrograph of interelectrode short-cicuiting bridge of PECVD-Si_3N_4 passivated device. Angle of tilt is 75^0

EFFECTS OF HIGH CURRENT AND TEMPERATURE IN POWER MESFET METALLIZATIONS

Claudio CANALI (:), Fabio CHIUSSI (::), Leonardo UMENA (::), Massimo VANZI (::), Enrico ZANONI (:::)

(:) Dipartimento di Elettronica ed Informatica, Universita' di Padova, Via Gradenigo 6/A, 35131 Padova (Italy)
(::) Telettra S.p.A., Reliability and Quality Department, Via Trento 30, 20059 Vimercate (MI) (Italy)
(:::) Dipartimento di Elettrotecnica ed Elettronica, Universita' di Bari, Via Re David 200, 70125 Bari.

Effects of high current density and temperature closely combine to degrade power MESFETs during their operating life in radio-link systems. To understand failure mechanisms and distinguish between those accelerated by high current and/or by temperature, we have performed various dc tests and measured thermal resistance and thermal maps of tested devices.

1. INTRODUCTION

The reliability of microwave communication systems markedly depends on the long term stability of GaAs power MESFET's. Degradation of these devices has been observed both in field applications and after rf or dc accelerated tests.

In real applications, burn-out is the dominant failure mode; however, owing to its catastrophic nature, the phenomenon has not been clearly understood. Rf accelerated tests are closer to real operating conditions of devices, but do not enable the effects of different acceleration factors and in particular of high current density and high temperature to be evidentiated.

In the framework of a large program aiming at evaluating the reliability of commercially available power MESFET's of different technologies and suppliers, we ran various dc accelerated tests over 10000 hours and measured thermal resistance and thermal maps of tested devices. Part of the results obtained were previously reported (1,2). Here the attention is focused on the results of dc accelerated tests which enabled us to distinguish between failure mechanisms accelerated by high current density or by high temperature and to understand degradation effects observed when these two accelerating factors are closely combined.

To this end, different dc tests were run, including:

i) High Forward Gate Current (HFGC) test at T_{case} = 200 °C and j = 5 10^5 A/cm^2 through the section at the beginning of each gate finger which was forward biased. Temperature increase of the device due to power dissipation during this test was very limited (3 °C), so that T_{ch} = T_{case} = 200 °C. This test enabled us to evidentiate electromigration effects in gate fingers.

ii) High Temperature Storage (HTS) test at T_{case}=250 °C without bias, in order to evidentiate failure mechanisms accelerated by temperature only.

iii) DC Operating Life (DCOL) tests at different channel temperatures up to 250 °C. In these tests devices were biased at the operating point, so that the increase of temperature due to the dissipated power was significant, and, as a consequence, high current and high temperature were combined as accelerating factors. In order to understand the results of these tests, thermal resistance and thermal maps of devices were measured.

2. FAILURE MECHANISMS INDUCED BY HIGH CURRENT DENSITY

2.1 Electromigration of Al-gate fingers

A gate current density of 5 10^5 A/cm^2 caused electromigration in Al-gated devices even after only 1000 hours of HFGC test. As shown in Fig. 1, the Al finger interruption can be clearly observed with both secondary electron SEM (Fig. 1a) and electron-beam-induced current (EBIC) techniques (Fig. 1b). Most interruptions occur at the beginning of the fingers, where the current density is higher (1) and the metallization could be thinner owing to the mesa step. From the electrical point of view this phenomenon prevents pinch-off being achieved and increases the gate series resistance.



Fig. 1 a) SEM image of Al-gate finger breakage due to electromigration, after 1000 h of HFGC test; b) EBIC image of same finger.

2.2 Al/GaAs interaction with increase of barrier height

Al-gate/GaAs interaction was observed during HFGC tests and causes an increase of the barrier height from 0.80 +/- 0.05 eV to 0.97 +/- 0.01 eV even after only 24 hours, Fig. 2. This effect does not occur in pure thermal tests (1) and appears to be due to the formation of an $Al_xGa_{1-x}As$ at the interface. The growth of this layer is markedly enhanced by high current density through the junction, 5.4 10^3 A/cm^2 for the HFGC tested sample of Fig. 2.

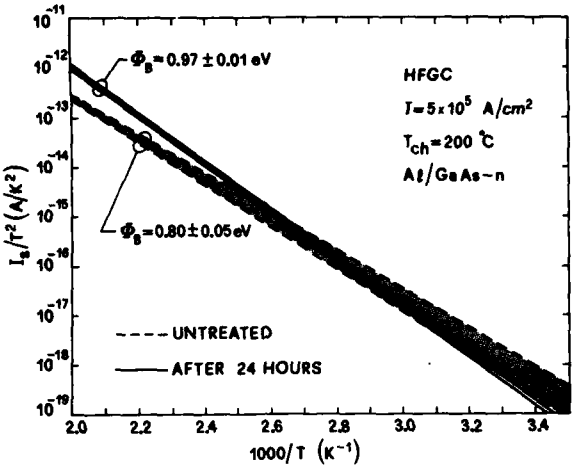


Fig. 2 Activation energy plot for the measurement of barrier height of Al-gate MESFETs before and after 24 h of HFGC test.

2.3 Ohmic contact electromigration

High j through ohmic metallizations induces Au electromigration at the drain contacts, Fig. 3. The SEM images of Fig. 3b and c) refer to drain contacts of the device after 5000 hrs. of DCOL test with T_{ch} = 200 °C and j = 5.3 10^5 A/cm^2 across drain finger. Gold removal at the end and accumulation at the beginning of the drain finger are evident, in agreement with electron wind.

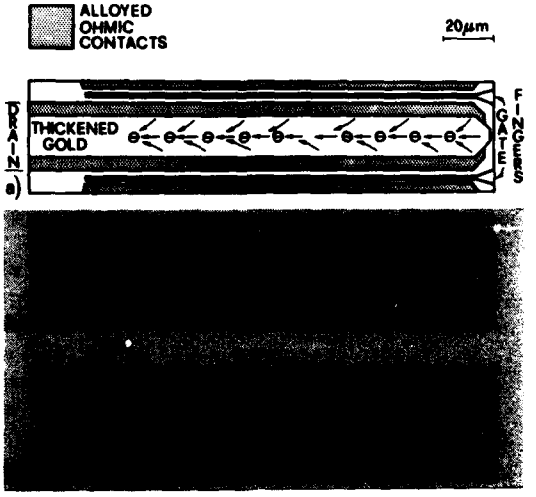


Fig. 3 a) Schematic diagram of electron wind along drain finger; b), c) SEM images of a drain contact on a device that endured 5000 h of life testing at T_{ch} = 200 °C and J = 5.3 10^5 A/cm^2.

3. FAILURE MECHANISMS INDUCED BY THERMAL STRESS

In devices with gate metallization based on Au (Ti/Au, Ti/W/Au, Ti/Pd/Au , Ti/Pt/Au, etc.), major reliability problems arise owing to Au interdiffusion through barrier layers. This process is mainly thermally activated and leads to a gate metallization "sinking" into the underlying GaAs, causing a reduction of effective channel height, with a consequent increase of the open channel resistance R_o and a decrease of I_{dss} and of pinch-off voltage V_p. We observed this failure mechanism in all devices stressed in HTS test (2).

4. ANALYSIS OF THERMAL INHOMOGENEITIES

DC Operating Life test as well as rf tests and in-field operating conditions do not allow an easy understanding of degradation induced by current density or by temperature alone. In fact, when the device is biased in the above mentioned conditions, thermal gradients may arise, due to local differences in the heat dissipation.
Local inhomogeneities, both thermal and structural, strongly affect the hardness of devices against either electromigration and interdiffusion and are of dramatic importance in GaAs MESFET's owing to their small geometries. In fact, in devices which failed in DCOL tests, not only is degradation faster, but also failure sites are differently distributed.

The evaluation of the channel temperature T_{ch}, including the effect of power dissipation is usually performed by means of thermal resistance measurements employing ΔV_{gs} method (3). However, owing to the modular structure of the devices and the possible presence of defects, the junction temperature across the chip active area is inherently non-uniform. As a consequence, it is not clear whether the T_{ch} values obtained by means of the R_{th} method refer to the average temperature on the mesa area or on the gates area, or include the possible presence of hot spots along the gate fingers.

Thermal behaviour of biased devices was then analyzed by means of a high resolution IR microscope, Barnes CompuTherm, which allows a temperature sensitivity better than 0.5 °C, a spatial resolution = 15 μm and the automatic correction of surface emissivity; the microscope provides digital maps (128x128 pixels) and line profiles, and calculates average values of temperature on arbitrarily defined areas of the chip.

Results were compared with T_{max} values obtained by R_{th} measured with the V_{gs} method. The main features are:

a) In devices with extremely low values of R_{th}, few °C/W, and with excellent thermal uniformity, the average temperatures of the mesa and of the gate areas differ only by few °C and are approximately equal to T_{ch}.

b) In some devices which show higher values of R_{th}, marked differences appear between the average temperature of chip, mesa and gate areas, and strong local thermal gradients along the gate finger as reported in Table I. Furthermore, in these devices $T_{ch}(\Delta V_{gs})$ is higher than the average temperature of mesa and gates area and approaches the maximum values obtained along the gate fingers.

TABLE I

R_{th} (measured by ΔV_{gs} method) = 49.2 °C/W
Dissipated power P_d = 980 mW
T_{case} = 31.7 °C Four gates device
$T_{ch} = T_{case} + R_{th} P_d$ = 79.9 °C

ZONE	$T_{average}$ (°C)	T_{max} (°C)
CHIP	45.9	
MESA	48.6	
GATES 1-2	65.5	
GATES 3-4	62.6	
GATE 1	73.2	81.4
GATE 2	76.6	85.6
GATE 3	73.7	80.4
GATE 4	73.3	81.0

Figs. 4 and 5 report the temperature profiles measured by means of IR microscope perpendicularly and across gate finger in one of these devices. Temperature profiles taken along gate width show that the hottest area is located at the beginning of gate finger, where the current density and the power dissipation is higher. This explains why electromigration and Al-gate interruption normally occurs at this site.

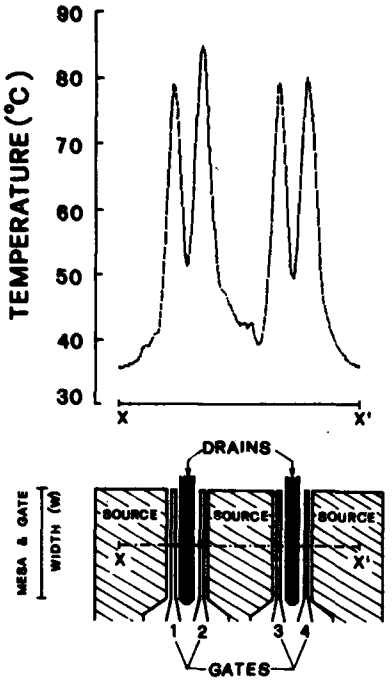
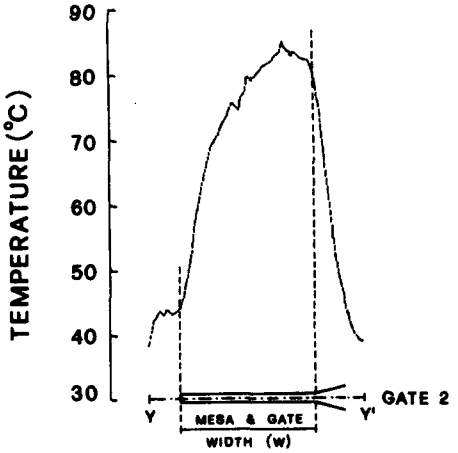


Fig. 4 Temperature profile measured perpendicularly to the gate fingers.



★ GATE LENGTH-WIDTH RATIO NOT IN SCALE

Fig. 5 Temperature profile measured along a gate finger.

In some devices, and in particular when the high values of R_{th} arise from non-uniform die attach and/or chip thickness, the highest temperature values can be found even in the central zone of the gate. In this case, consequently, metal/GaAs interaction results enhanced in the central area of the gate, as shown in Fig. 6 for a Ti/Au gated MESFET after 4000 hrs. of DCLT at T_{ch} = 200 °C. The three magnified SEM micrographs of the rear side of the gate contact, after the removal of GaAs, reported in Fig. 6b, clearly show that the strongest interaction occurred in the central zone of gate finger.



Fig. 6 SEM micrograph of gate finger of a MESFET after 4000 h DCOL test at T_{ch}=200 °C. a) Rear side. b) Magnified zones.

5. CONCLUSIONS

Degradation mechanisms induced by high current density (electromigration in Al-gate fingers, Al/GaAs interaction with a consequent increase in barrier height, Au drain electromigration) and by temperature only (Au metal gate/GaAs interdiffusion) were identified by means of dc accelerated tests.

In real applications both accelerating factors, T and j, are combined, and local thermal inhomogeneities markedly influence the device failure mechanisms.

High resolution IR microscopy enables to understand degradation phenomena and to verify that in devices with high R_{th} values the calculated $T_{ch}(V_{gs})$ value approaches the highest temperature measured along gate width.

Work partially supported by CNR, P.F. MADESS

REFERENCES

(1) C.Canali, F.Fantini, A.Scorzoni, L.Umena, E.Zanoni, IEEE Trans. on Electron Devices, Vol.ED-34 (2), p.205 (1987).

(2) C.Canali, F.Castaldo, F.Fantini, D.Ogliari, L.Umena, E.Zanoni, IEEE El. Dev. Letters, Vol. EDL-7(3), p. 185 (1986).

(3) H. Fukui, IEEE Int. Electron Device Meeting 1980, p. 118.

FAILURE MECHANISMS STUDY OF A STANDARD GaAs IC TECHNOLOGY

G. KERVARREC, J.M. DUMAS, J.Y. BOULAIRE and J.F. BRESSE*

Centre National d'Etudes des Télécommunications - BP 40 - 22301 LANNION - FRANCE
*Centre National d'Etudes des Télécommunications - 196, Av. Henri Ravera - 92220 BAGNEUX - FRANCE

After more than ten years of research and development in many laboratories, GaAs logic integrated circuits are becoming available for high speed logic applications. Reliability is one of the requirements needed for their introduction in systems and this is the purpose of this contribution.

1. INTRODUCTION

In order to find the degradation mechanisms and establish the reliability of such devices, we have carried-out life tests on drop-in chips processed on the same wafers as high frequency dividers. (The basic logic gate configuration is the Buffered FET logic (BFL)).

A drop-in chip contains various test-patterns (FETs, diodes, site-gate pattern, isolation and metallisation patterns, etc...) developped for the process control and also to study the parasitic effects.

100 chips issued from three wafers have been mounted in dual-in-line ceramic packages (DIL 24), the pattern bond pads being connected to the pins. Such packaging allowed us to conduct biased life tests appropriate to each test-pattern. A set of thirty-five DC parameters was measured on the drop-in chips over the whole aging experiment with an automatic test equipement.

No major degradation has been observed on the metallisation patterns, ohmic contacts and Schottky gate. This means that the improvements made on the metallurgy of GaAs discrete devices have been successfully transferred to the GaAs IC technology. The isolation provided by the 5000 Å thick PECVD Si_3N_4 layer between the two TiPtAu metallisation levels, did not change during aging. A slight increase has been measured on the side-gating threshold voltage.

However degradations have been identified and investigated on the FETs.

2. FAILURE MECHANISMS STUDY

The degradations can be separated into two types :

(a) Some FETs presented, at the origin, high gate-to-drain and/or gate-to-source leakage currents and they have not been driven in life-tests.

(b) The others FETs (with no leakage) have been aged. Short circuits, mainly between gate and drain, occured under biased life-tests, indicating the degradation is elecric-field induced.

Aged and not-aged devices have been then analyzed and the following results can be reported :

1 - Leakages and shorts were due to conductive paths (maximum width ≈ 0.5 µm) developping between the gate and ohmics contacts, at the GaAs surface, in the access regions.

2 - These paths, which are invisible using the classical SEM observation and the associated EBIC mode, have been identified using Scanning Optical Microscopy (SOM) [1] and the associated Optical Beam Induced Current mode (OBIC). Such investigation is non-destructive for GaAs. Moreover, as the optical absorption in a semiconductor is related to the wavelength (λ), the paths have been located at a maximum depth of ≈ 500 Å into GaAs by using λ = 442 nm (HeCd laser), (see Fig. 1).

3 - High resolution Auger analysis [2] (spot size ≈ 1000 - 2000 Å) have been then conducted (see Fig. 2) after "in-situ" etching of the Si_3N_4 layers. Spectrum (1) has been taken at the GaAs-

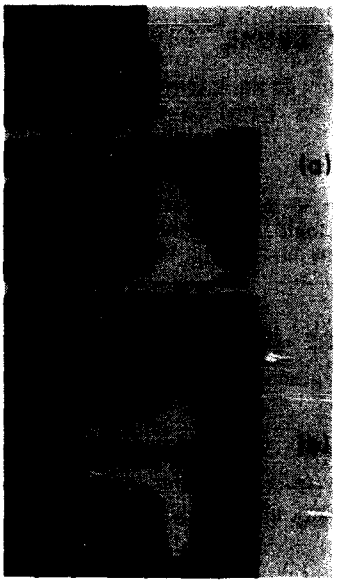


FIGURE 1
SOM views of a "path" between gate and drain : (a) reflected mode ; (b) OBIC mode.

RF sputtered Si_3N_4 interface, (this Si_3N_4 layer was used as implantation cap and selectively open for the gate recess), far from the active area : it shows high amounts of oxygen and arsenic. These amounts strongly increase when the analysis is driven on a "path" existing on FETs at the origin, i.e. before aging (spectrum (2a)). After aging, the signature of gold (69 eV) is also found.

4 - Presence of oxygen, at the GaAs-RF sputtered Si_3N_4 interface, has been confirmed using Secondary Ion Mass Spectrometry (SIMS), (Fig. 3). Sulfur (S) is also clearly detected at the same interface. Regarding the results of physical analysis, the following mechanisms could be involved : (i) at the origin : the path can be created by the gradual release of As due to a GaAs oxidation mechanism during processing :

$$As_2O_3 + 2GaAs \rightarrow Ga_2O_3 + 4\ As \quad [3]$$

and the conduction made through a thin layer of elemental metallic As, [4], (ii) during aging : an electric-field induced migration of electrode material (Au from the ohmic contacts) - depending strongly of the presence of As oxide (surface not perfectly treated with sulfurious acid solution,

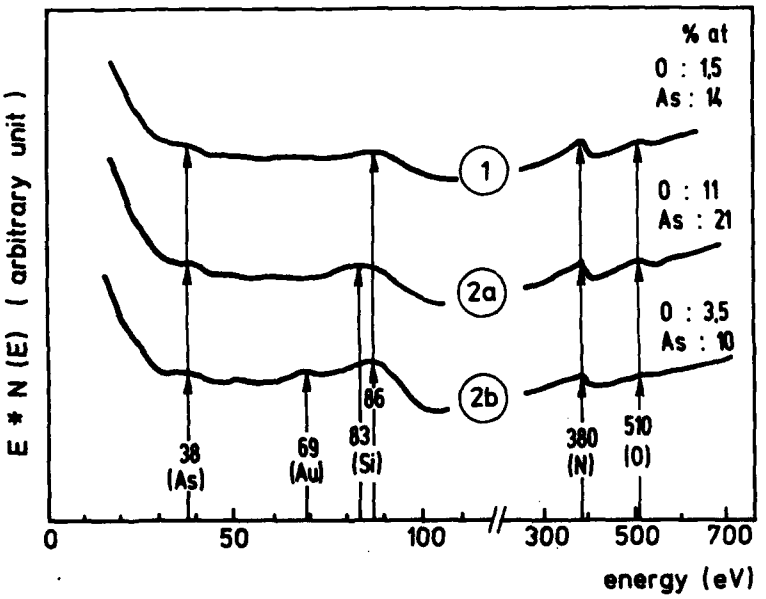


FIGURE 2

Auger analysis : (1) far from the active area ; (2a) on a "path at the origin ; (2b) on a "path" after aging.

as suggested by the presence of sulfur - Fig. 3), [5], created the path.

3. THE GATE-LAG EFFECT

The gate-lag effect corresponds to a drain-current transient in response to a voltage pulse on the gate with little drain-to-source voltage (V_{DS}) variation. Although several research groups are studying this parasitic effect limiting the performances of GaAs devices, few results have been published until now [6], [7]. Fig. 4 shows the electrical schematic of the measurements we have performed on the FETs. The gate is switched form the pinch-off voltage to V_{GS}=0. The pulse width can vary in the range 1 µs to several seconds with 10 % repetition rate. The drain current response is measured through Rc.

Fig. 5a represents the rising edge part of the gate pulse. Fig. 5b is the corresponding drain current response before aging : the static value of I_{DSS} (7 mA) is instantaneously reached. After aging (Fig. 5c), the drain current switches partially on, from I_{DS} = 0 to I_{DS} ≃ 4 mA, and then lags to the static valve of I_{DSS} (7 mA) with a time constant > 100ns. It is obvious that 4 mA is the maximum drain current available for microwave operation. Such phenomenon can affect the high frequency operation of an integrated circuit, when the gate switches capacitively. Surface effects are involved in this degradation mechanism [6], [7], but they are not well understood in the present case.

4. CONCLUSION

Failure mechanisms have been investigated on a standard GaAs IC technology by means of physical analysis. Successfull technological improvements have been then conducted.

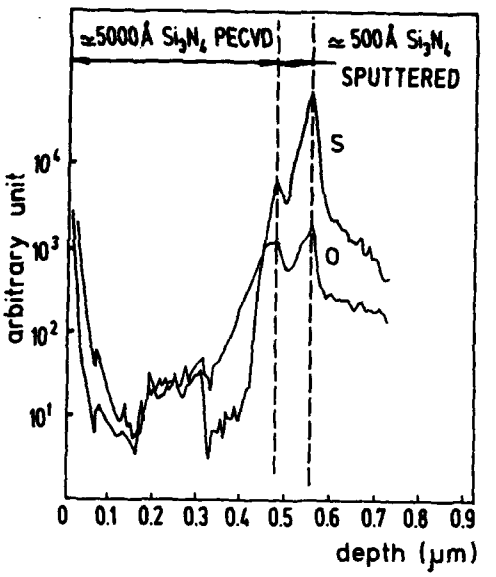


FIGURE 3
SIMS profiles of oxygen and sulfur

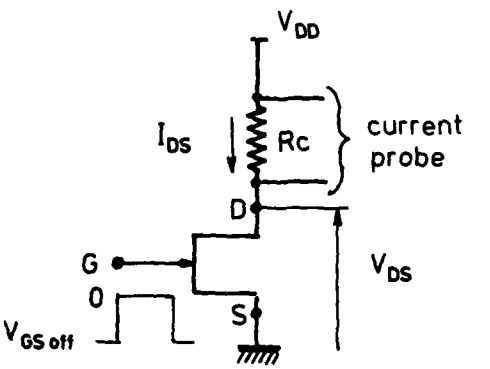


FIGURE 4
Electrical schematic diagram of gate-lag measurement.

REFERENCES

[1] Wilson, T., and Sheppard, C., Theory and practice of scanning optical microscopy, (Academic Press Ltd, London, 1984).

[2] Mogami, A., Thin Solid Films, vol. 57, (1979), 127.

[3] Thurmond, C.D., Schwartz, G.P., Kammlott, G.W., and Schwartz B., J. Electron. Soc., Vol. 126, N°6, (1980), 1366.

[4] Folkes, P.A., Chang, C.C., and Lane, E., J. Electrom. Soc., Vol. 132, N°6, (1985), 1417.

[5] Kretschmer, K.H., and Hartnagel, H.L., 23 rd Ann. Proc. Int. Reliab. Physics Symp., Orlando (Fl) (1985).

[6] Rocchi, M., Proceedings of European Solid State Devices Research Conference, (North-Holland, Amsterdam, 1985).

[7] Dumas, J.M., Garat, F., and Lecrosnier, D., Elect. Lett., Vol. 23, N°4, (1987), 139.

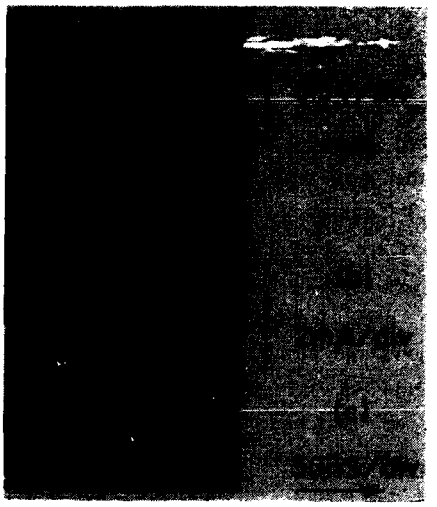


FIGURE 5
Gate-lag measurements : (a) rising edge part of the gate pulse ; (b) drain current response before aging, and (c) after aging.

RELIABILITY OF LOW-NOISE MICROWAVE HEMT BY MOCVD

K. Tanaka, H. Takakuwa, J. Kobayashi and Y. Kato

Sony Corp. Semiconductor Group
4-14-1, Asahicho, Atsugi,
Kanagawa 243, JAPAN

Low-noise HEMTs were evaluated for reliability by conducting high temperature accelerated life tests and examining the changes in electrical characteristics of the device as well as visual signs of degradation. The reliability of the HEMTs was found to be on the same level as that of conventional GaAs MESFETs, and no degradation associated with the unique structure of the HEMT was observed.

1. INTRODUCTION

HEMTs (High Electron Mobility Transistors) for low-noise microwave applications have been introduced by various manufacturers and are expected to eventually replace the conventional GaAs MESFETs in small-signal front-end applications.

The current HEMTs employing AlGaAs/GaAs heterojunction structures have demonstrated performance superior to the best available MESFETs today, with minimum noise figures well under 1.0 dB at 12GHz and 1.5 dB at 18 GHz [1],[2].

HEMTs have a more complicated structure than conventional MESFETs, and this paper describes the results of high temperature storage tests. The data described here include performance of actual HEMTs (200-micron wide, 0.5-micron long gates) under high temperature storage (at 200°C, up to 2000 hours), as well as observations of various physical parameters obtained from process monitor elements heated to 300°C.

2. DEVICE FABRICATION

The cross-sectional structure of a typical low-noise microwave HEMT with AlGaAs/GaAs heterojunctions and AlAs/GaAs underlying super lattice structures is depicted in Fig. 1. A photograph of a 200-micron wide chip is shown in Figure 2.

The epitaxial growth necessary for the formation of the HEMT heterojunction and superlattice is performed by MOCVD using trimethyl aluminum (TMA), trymethyl gallium (TMG) and arsine carrier gas (AsH_3) under atmospheric pressure. The growth temperature is 720°C and the growth rate is about 240Å/min. The operation of the reactor is managed by an automated sequence controller, resulting in a high level of reproducibility and uniformity of the epitaxial growth process.

Hall mobilities of the two dimensional electron gas of the actual device at the interface are 8000 and 30000 cm2/V-sec at 300 and 77 K, respectively, with an undoped AlGaAs

spacer layer of 10 Å. The mobility and the sheet carrier concentration of the two dimensional electron gas are comparable to those reported using MBE. A high mobility and a low sheet resistivity of the two dimensional electron gas are the two most important parameters for the realization of low noise HEMTs. A thin super lattice buffer of alternating undoped AlAs and GaAs layers and a 10 Å-thick undoped AlGaAs spacer layer were introduced to satisfy the above requirements as shown in the device cross section.

The aluminum gate metal is evaporated at an angle to offset the actual Schottky contact region towards the source and partly overlap onto the spacer sidewall. This results in a reduction of the effective gate length together with a decrease in the series gate resistance resulting from the high cross-section-to gate-length ratio. The Au-Ge/Ni ohmic metal for source and drain is defined by standard lift-off using silicon nitride as the lift-off spacer. All currently fabricated HEMTs are passivated with a 500 Å layer of silicon nitride after the definition of the gate and ohmic regions.

3. HIGH TEMPERATURE LIFE TEST

For the fabrication of HEMTs, a high purity GaAs layer and an abrupt AlGaAs/GaAs interface are required. In addition, the uniformity and quality of the epitaxial layers must be well established. To date, for the fabrication of semiconductor laser diodes which require very high reliability to assure a long lifetime, only LPE and MOCVD have been successfully applied.

Figure 3 shows the results of high-temperature (200°C) storage tests in terms of the device's RF performance (noise figure and associated gain). It can be seen that no degradation of either parameter occurs at 200°C even at the end of the 2000-hour test period. DC parameters were also checked and found to be unaffected.

Once it became clear that there is no measurable degradation at 200°C storage, additional tests of DC parameters were performed at 300°C using process monitor elements.

Characteristics of the Schottky gate such as the capacitance of the built-in junction (Co), n-value and forward voltage (V_f) do not show any change after a 300°C, 150-hour storage, as shown in Fig. 4. The aluminum Schottky barrier to the underlying n-doped AlGaAs is thus shown to be a stable gate structure for the fabrication of HEMTs.

DC parameters of the HEMTs such as pinch-off voltage (Vp), transconductance (Gm), saturated drain current (Idss) are also unchanged after the 300°C, 150-hour test, as shown in Figure 5. These data indicate that contrary to some reports that DC parameters, particularly Idss, show fluctuations and require a stabilization bake, the electrical parameters and the physical parameters of the HEMT are inherently stable.

The sheet resistivity of the two dimensional electron gas at the interface of the heterojunction does not change after the 300°C, 150-hour storage, as shown in Figure 6. This indicates that any interdiffusion

between Si, Ga, As and Al atoms at the interface does not give adverse effects to the 2DEG channel.

A more severe test of the stability of the heterojunction was done by measuring the mobililty (u) and the sheet carrier density (ns) were after 30-minutes of heat treatment at 500, 600 and 700°C. They are shown in Fig. 7. At 700°C the mobility decreases and the carrier density increases, suggesting possible Si diffusion from the Si-doped AlGaAs region to the undoped GaAs layer. However u and ns remain constant up to 600°C. At the end of the 150-hour test, the ohmic contact resistance however does show a gradual increase indicating a deterioration of the ohmic alloyed region, as shown in Fig. 8. This data is similar to that observed from standard GaAs MESFET ohmic contacts which also use the AuGe/Ni eutectic alloying procedure, and the cause is assumed to be equivalent for both HEMTs and MESFETs. An activation energy of approximately 1.9eV has been derived by an Arrhenius plot of the ohmic contact resistance. Using this value, the extrapolated MTTF of the HEMT at room temperature due to ohmic degradation (defined as the time it takes for the ohmic contact resistance to double) would be on the order of 10^7 to 10^8 hours.

The results so far show that HEMTs fabricated by MOCVD are very reliable in addition to offering the best noise performance. The current HEMT devices are as reliable as MESFETs, and no degradation associated with the AlGaAs/GaAs heterojunction or the AlGaAs/metal interface was observed. More comprehensive data on the various failure modes, particularly those related to the ohmic metal system, the heterojunction interface and the superlattice will be collected and reported in the future.

REFERENCES

[1] K. Tanaka, M. Ogawa, K. Togashi, H. Takakuwa, H. Ohke, M. Kanazawa, Y. Kato and S. Watanabe, "Low-Noise HEMT using MOCVD," IEEE Trans. on Microwave Theory and Techniques, Vol MTT-34, No. 12, Dec. 1986, pp. 1522-1527.

[2] H. Takakuwa, K. Tanaka, Y. Mori, M. Arai, Y. Kato and S. Watanabe, "A Low-Noise Microwave HEMT using MOCVD," IEEE Trans. on Electron Devices, Vol. ED-33, No. 5, May 1986, pp 595-600.

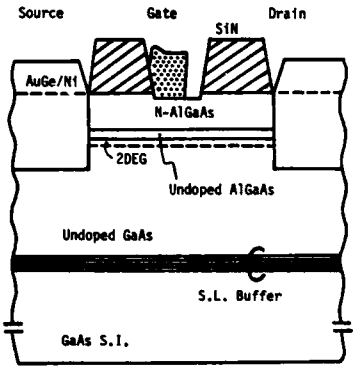


Fig. 1 Cross section



Fig. 2 Photograph (Wg=200um, Lg=0.5um)

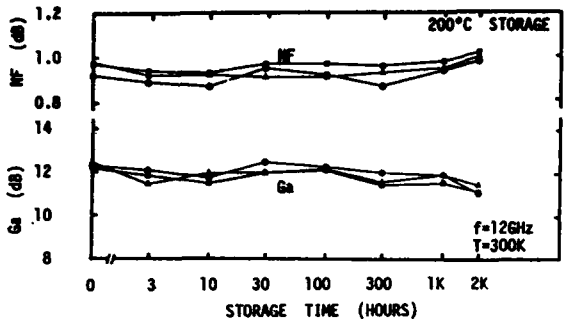


Fig. 3 NF,Ga after 200°C storage up to 2000H

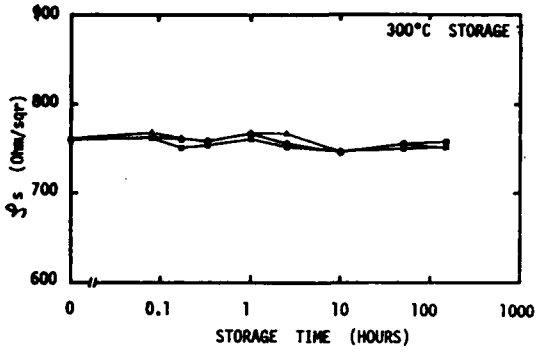


Fig. 6 Changes in ρ_s after 300°C storage

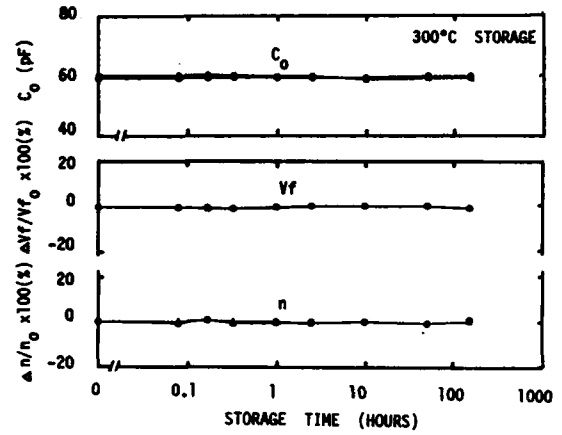


Fig. 4 Changes in Co,n, and Vf after 300°C storage

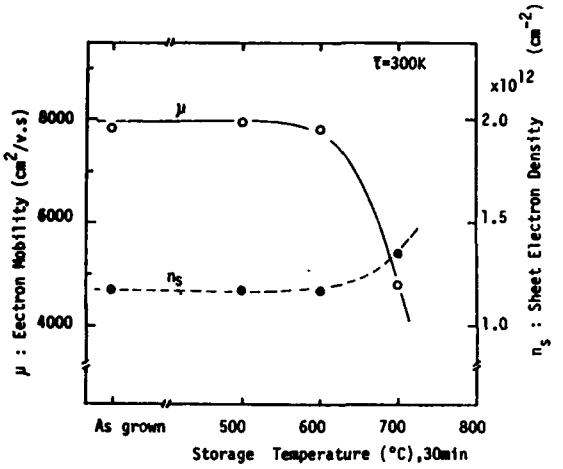


Fig. 7 μ and n_s as a function of temperature

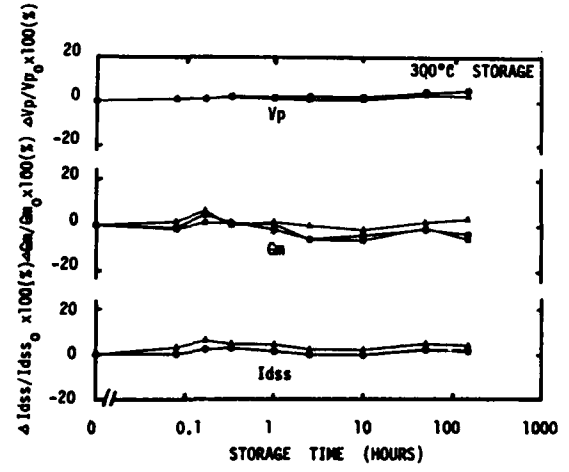


Fig. 5 Changes in Vp,Gm,Idss after 300°C storage

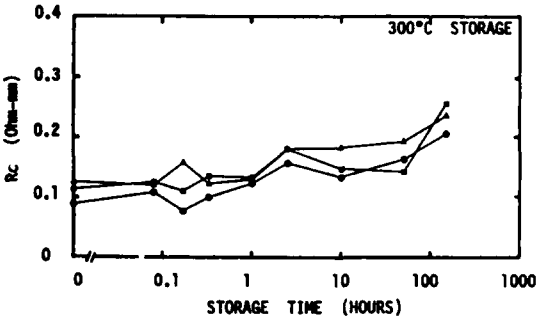


Fig. 8 Changes in Rc after 300°C storage

ION BEAM LITHOGRAPHY

Gerhard Stengl, Hans Löschner, Ernst Hammel,
IMS - Ion Microfabrication Systems GmbH,
A-1020 Vienna, Austria

Edward D. Wolf,
National Nanofabrication Facility,
Cornell University, Ithaca, New York 14853, USA

1. INTRODUCTION

With great effort advanced lithographic techniques are being developed to cope with immediate production needs of devices with submicrometer design rules. Optical 1:1 and reduction techniques will fulfill the lithographic requirements of pattern transfer down to 0.5 µm.

This paper is focused on the potential of ion beam lithography for the development, customization and high volume production of sub-0.5-µm devices.

Ion beam lithography (IBL) is being developed in three directions (Fig. 1): Focused (FIBL), Masked (MIBL, using channeling or open stencil masks), and Ion Projection Lithography (IPL, using open stencil masks).

2. FOCUSED ION BEAM LITHOGRAPHY (FIBL)

Using high brightness liquid metal ion sources FIB ion-optical columns can realize spot sizes down to 0.03 µm with current densities up to 10 A/cm² [1]. FIB lithography for quantum wells with 0.06 µm diameter on 0.5 µm centers in arrays of 1024 x 1024 dots (500 x 500 µm², was done in 18 sec [2].

Recently, considerable attention has been given to the development of high brightness hydrogen and helium gaseous field ion sources [3]. Light ions deposit their energy in small, well defined volumes with penetrations depths of about 1 (0.7) µm for 100 keV H^+ (He^+) ions [4].

But all these sources deliver a total usable current of just several 100 pA and provide $\leq 10^8$ ions/sec on the substrate. Thus, pixel transfer rates are limited to 10^7 Hz (with 0.05 µm pixel size and about 100 ions exposed into one pixel site [5]).

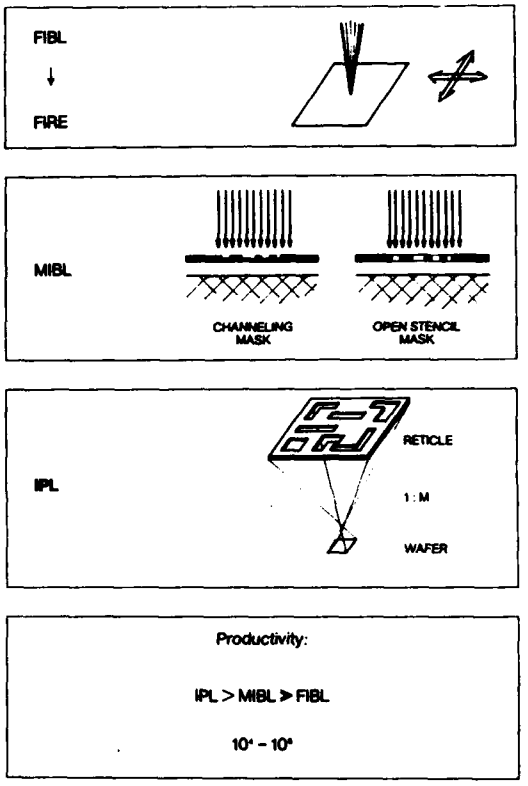


FIGURE 1. Ion beam lithography development.

Furthermore, compared with scanning electron beam lithography, the slow velocity of the accelerated ions imposes severe limitations in the writing speed. Traveling wave deflectors may speed up the pixel transfer rates [6,7] but will complicate FIB systems considerably and will add a substantial reduction of chromatic errors to the ion-optical column.

Consequently, with respect to deflection speed, the best choice for lithographic applications is the use of ≥ 100 keV H^+-ions and the realization of variable shaped beams.

Compared with the well established E-beam tools, a scanning FIB system would ensure much less proximity effects leading to more simple pattern generation. This may be important for the realization of photomasks and reticles for x-ray lithography with submicron feature sizes [8]. But mask making technologies may move towards bi- or trilevel resist schemes which overcome the disadvantage of electron beam scattering effects. These effects are also substantially reduced by using ≥ 50 keV electron energies.

Because of these reasons, the main impact of H^+-FIB lithography may be in the field of prototyping of new devices where the lack of proximity effects gives an important advantage in process technologies. A FIB system may provide not only lithographic pattern generation but also the possibility of ion-assisted etching [8] and deposition [9] and localized sputter removal of surface layers. The same system can be used as an in situ ion microscopy and SIMS analyzing tool [10]. (These features are already used for industrial mask repair systems [11].)

For submicron device prototyping and customization the focused ion beam must fulfill requirements on positional accuracy of the order of 0.01 µm. Such precision could be achieved for FIB systems using principles of demagnifying ion projection optics [12,13].
It should be mentioned that - using a source delivering ion currents of several 100 µA [14] - the inherent limitations in single beam scanning could be solved by multiple beam exposure techniques either using switchable beamlet apertures [15] or screen lens projection [16]. Severe problems of positional accuracies and uniformities of beamlet currents have to be overcome in this case.

3. MASKED ION BEAM LITHOGRAPHY (MIBL)

As with x-ray lithography the MIBL technique uses thin Si (or Si_3N_4) membranes as mask reticle support. In the "all Si" approach the absorber is formed by the thicker (2 µm) regions of the membrane whereas the incidenting 100 keV - 200 keV H^+ or He^+ ions are channeled through the thinner (0.5 µm) support areas [17]. The channeled ions leave the mask with a scattering angle of about 4°. Thus theoretical and experimental studies show resolution limits of 0.5 µm pattern transfer into single layer organic resists [18].

Implementing open stencil masks, ion beam 1:1 exposure demonstrates sub-0.1-µm resolution in PMMA organic resists with excellent process latitudes [19]. Covering the openings with a fine line grid, complicated patterns were realized on the wafer (rocking the ion beam during exposure). Another solution of the "hole in the doughnut" problem is the splitting of a lithographic design into a set of complimentary open stencil masks. So far, this technique has been developed for 1:1 electron beam shadow printing [20].

To realize the required control of mask pattern distortion the openings may only be realized in mask foils under low (≈ 10^9 dyn/cm²) stress [21]. For channeling or open stencil MIBL techniques mask heating during chip exposure is setting a limit to the power load of about 50 mW/cm² of the ion beam impinging on the mask foil [21]. For 100 keV ions thus dose rates of several 10^{13} ions/cm²/sec can be realized. Statistical requirements for reliable sub-0.5-µm pattern transfer ask for a resist with a sensitivity of ≥ 10^{12} ions/cm² [6,7] which limits the (0.05 µm) pixel transfer rates and thus MIBL productivity to the order of 10^{11} Hz [5].

4. ION PROJECTION LITHOGRAPHY (IPL)

The schematics of an Ion Projection Lithography Machine (IPLM) are shown in Fig. 2. A duoplasmatron is used as the ion source [14] where depending on the ion species (H^+, He^+, N^+, Ar^+) up to 200 µA can be extracted with angular current densities of > 100 mA/sr (to be compared with < 100 µA/sr for liquid metal ion sources). The ions seem to originate from a virtual source whose size is only ≈ 5 µm Ø [22]. The extracted ions are focused to an ExB filter region. Subsequently a broad ion beam illuminates the open stencil mask reticle with ion energies between 5 keV and 10 keV. The ions passing through the mask openings enter an immersion gap lens which accelerates the ions to their final energy of 50 keV - 100 keV (or higher, depending on system design). The immersion lens focuses the ion source into the projective lens system. This Einzel lens projects an ion image of the mask opening patterns at a defined reduction factor (3x-20x) onto the substrate. The desired ion species is selected through ExB action at the (variable, Fig. 2) entrance aperture of the projective lens.

Applying appropriate voltages to the rods of an octopole, placed in front of the projective lens, electrostatic dipole fields are generated which can move the projected ion image in the X-Y plane of the wafer target. A fine rotation of the ion image can be achieved by axial magnetic fields with a solenoid surrounding the octopole site. (A coarse rotation of the ion image is achieved by mechanical reticle movement.) A reference block between projective lens and wafer containes 8 channeltron detectors to obtain signals for die-by-die alignment. Chip exposure can be controlled by mechanical or electronic shutter arrangements. Wafer exposure is done with a stepper X-Y-Z-θ-stage.

4.1. IPL PROCESS LATITUDES AND DEPTH OF FOCUS

Results obtained with an experimental set up (IPLM-01, without ExB analyzing unit) have demonstrated resolution patterns of 0.1 µm lines and spaces. This resolution was obtained in PMMA organic resists with image reversal [22,23,24] techniques replicating open stencil mask openings of 500 lines/mm (1 µm mask opening width) at 10x reduction. The pattern transfer characteristics shows a linear relationship between the PMMA pattern linewidths and the width of the mask openings [24]. Figure 3 shows an IPLM-01 result of the dependence of several pattern linewidths vs. chip exposure time (ion dose). Using 80 keV He^+-ions the exposure was done with a 5x reduction projective lens in a chip field of 4 x4 mm^2. This considerable process latitude is in the same order as results of a 1:1 ion beam exposure (done with 50 keV H^+-ions with positive PMMA development [19]) as indicated in Fig. 3.

Because of the quasi-paraxial flow of the ions in the IPLM column a very large depth of focus is achieved. Changing the wafer Z-position by ± 4.5 millimeter the resolution is degraded only by 10 % [24]. Because of the divergent ion beam the scale of the projected ion image changes with the wafer Z-position. This change can be compensated by setting the voltage of the projective lens [23,24]. This electronic change of scale can be done within ± 0.5 % for 10 % resolution changes [24]. Thus, using a wafer stage with Z-adjustment, topographical structures on the wafer up to heights of 5 µm will not impose pattern displacements of more than 0.05 µm at the corners of chip fields up to 10 x 10 mm^2.

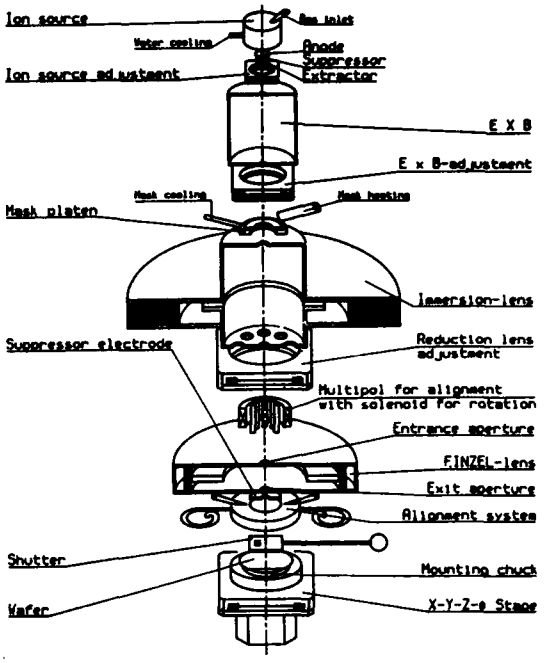


FIGURE 2. Principles of an Ion Projection Lithography Machine

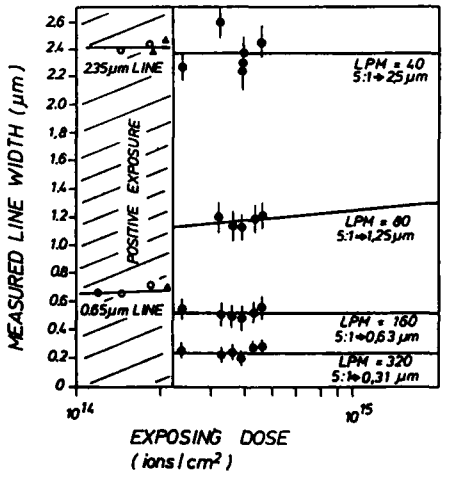


FIGURE 3. Results of linewidth measurements in PMMA organic resist vs. exposure dose. Comparison of 1:1 exposure (Randall [19], H^+-ions and positive development) with IPLM-01 exposure at 5x reduction (He^+-ions and image reversal development). LPM = line pairs / mm in the mask openings.

4.2. IPL OPEN STENCIL MASKS: FABRICATION AND EXPOSURE CONDITIONS

For the IPLM-01 experiments test masks, fabricated by Ni electroplating techniques [24], have been used. By clamping the Ni mask foil to an invar steel frame at elevated temperature (70 °C) the ion beam induced mask heating causes only a partial relief of the thermally induced stress but no pattern displacements. The distortion level of such test masks has been evaluated for continuous ion beam exposure conditions [22]. Because of the crude clamping of the 2 μm thick mask foil a 8 μm shear distortion was found for a 20 mm x 20 mm grid field [22].

An excellent lifetime of the Ni test masks of > 700 hours was found for a 5 keV He^+ ion beam exposure of the mask with current densities of ≥ 1 μA/cm² [14].

Within the framework of x-ray lithography, mask technologies have been developed to fabricate thin Si (or Si_3N_4, SiC) foils in situ connected to a supporting Si-wafer ring frame [25]. These mask blanks can also be used to fabricate open stencil masks.

Because of the divergent ion beam impinging on the open stencil mask, anisotropic mask openings have to be corrected in width as indicated in Fig. 4a. This is not necessary using tapered mask openings (Fig. 4b). The 5 keV ions are stopped within a < 0.1 μm surface layer of the open stencil mask providing excellent mask contrast. Thus, no linewidth errors are caused in the projected ion image. Fabrication techniques for tapered mask openings are already being developed [26] which may also use appropriate processes of Si trench etching technologies [27].

By simulating the fabrication of openings in a stressed mask foil the induced distortion was calculated by Randall [28] for the following exposure conditions: 5:1 reduction ion-optics, mask area = 5x5 cm², image field = 10x10 mm², 5 keV ion energy at mask, current density at mask = 1 μA/cm². The fabrication induced maximum image distortion of the 10 mm chip field is ≤ 0.05 μm even for the "worst case" (only half of the mask contains openings) when a 2 μm thick foil with a low stress of 10^9 dynes/cm² is used (Fig. 5). (Low stress mask foils can, e.g, be realized by adding Ge to the B doping of the Si wafer surface [25,26].)

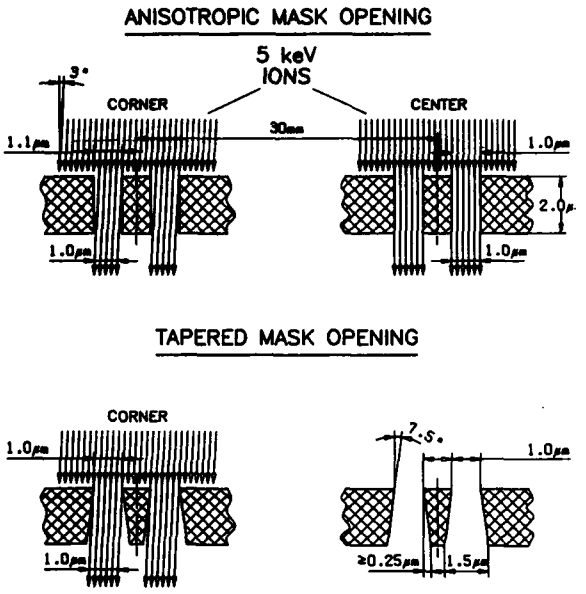


FIGURE 4. IPLM exposure conditions of open stencil masks.

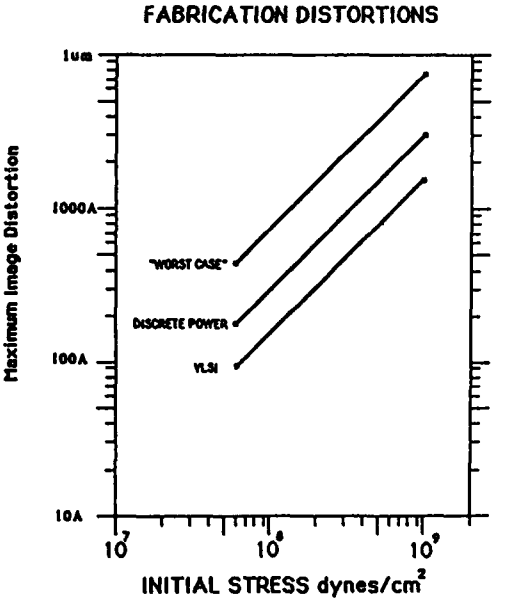


FIGURE 5. Maximum lateral fabrication distortion of projected mask image as a function of initial stress as calculated by Randall [28] for a 10 x 10 mm² chip field.

With the proposed use of such low stress Si open stencil masks, the IPLM induced heating would cause an expansion and thus distortions if the mask frame is pinned to room temperature. The amount of such exposure induced distortions is ≤ 0.05 µm for the 10 mm chip field for 5 mW/cm² mask exposure (Fig. 6, [28]).

Exposure induced distortions may be reduced to zero level [28] if the mask frame is heated to the temperature of the ion beam exposed mask foil [12]. In this way the total current extracted from the ion source can be used to expose the mask with current densities of 10 µA/cm² and power loads of 50 mW/cm².

Because of the immersion lens induced acceleration and the ion-optical reduction the power density at the wafer is enhanced by three orders of magnitude to several 10 mW/cm². For a 10x reduction system dose rates of several 10^{16} ions/cm²/sec may be realized. Thus, ion projection lithography may approach (0.05 µm) pixel transfer rates up to 10^{14} Hz which is equivalent to the productivity level of optical projection [5].

4.3. INSPECTION OF IPL OPEN STENCIL MASKS

With the goal to fabricate open stencil masks with mask openings down to 1 µm dimensions, commerical E-beam systems can be used for reticle pattern generation. Focused ion beam systems seem to be best suited for mask repair (with possible improvements using demagnifying ion projection principles [12] in the future).

But the currently used optical photomask inspection techniques are not sensitive to organic layers or particles in the mask openings which stop the impinging ion beam.

Because of these reasons a new mask inspection method using electron transmission microscopy has been developed. The principle is shown in Fig. 7. In an SEM the secondary electron detector is placed beneath the mask so that a transmission SEM picture of the mask openings can be obtained.

Figure 8a shows an example of a resolution test pattern with large dust particles as depicted in the conventional SEM micrograph of Fig. 8b.

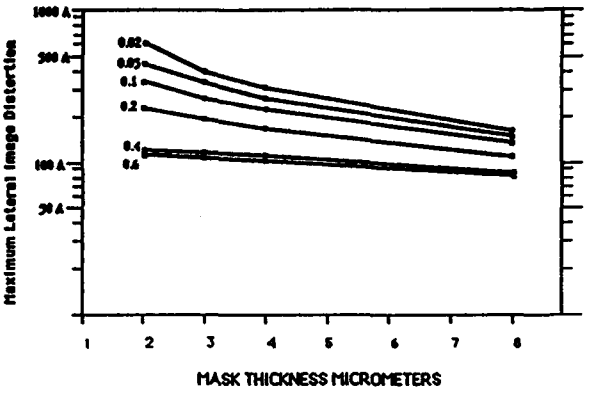


FIGURE 6. Distortions of the projected image (10 x 10 mm²) caused by mask heating in the cooled frame as a function of mask thickness with different values of emissivity as calculated by Randall [28].

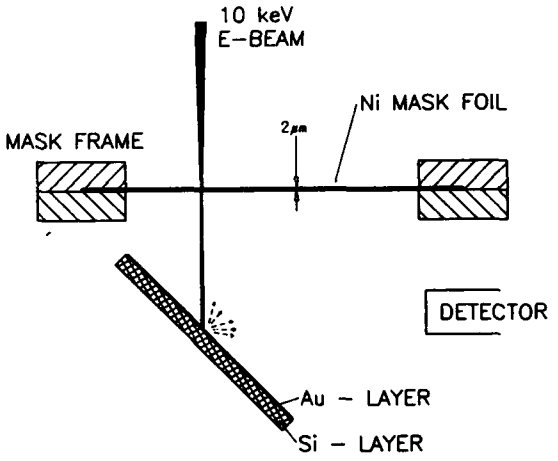


FIGURE 7. SEM transmission mode open stencil mask inspection.

In order to achieve the excellent transmission SEM contrast (Fig. 8a) electrons of 10 keV energy were used which are not penetrating the 2 µm thick Ni mask foil used for this experiment. The transmission SEM of the mask openings correlates with the ion image projected with the IPLM-01 onto the wafer (Fig. 9). This inspection method could also be implemented in an electron beam writing system so that defects and opening pattern positions of the fabricated stencil mask may be controlled in one step.

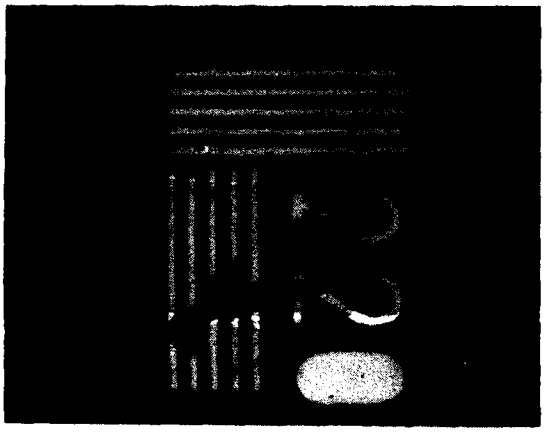


FIGURE 8a. SEM transmission mode photo of a Ni test mask opening with 220 lines/mm partly covered by defects.



FIGURE 8b. Conventional SEM photo of the defect covered test pattern.



FIGURE 9. Optical micrograph of an IPLM-01 image of the test pattern at 5x reduction (1100 and 1250 lines/mm, 0.91 and 0.80 μm pitch).

4.4. IPL PATTERN TRANSFER INTO ORGANIC AND INORGANIC RESIST LAYERS

The high IPL power densities allow the use of a wide variety of organic and inorganic resist materials [29].

Using organic (photo)resists and conventional (dip or spray) development the exposing ion beam must penetrate the resist material in order to obtain steep resist walls [4].

The damage induced by the ion beam in gate oxides, Si and GaAs substrates has been studied in context with annealing procedures. With stopping layer techniques any damage can be avoided [30] because the ion beam induced x-ray generation is negligible (less than 6 orders of magnitude in comparison with E-beam lithography).

Implementing dry development techniques, organic resists could be used for IPL pattern transfer as shown in Fig. 10.

Selecting ion species and energy the IPL exposure is done such that the ions stop within the organic resist or at layer sites where the induced defects do not degrade device performance. Then a transformation step (e.g., silylation as being developed for photolithography [31]) leads to a conformal mask for subsequent plasma development of the organic resist. (IPL exposure already in situ may lead to a selective surface transformation, e.g., the carbonization of organic resist materials [32,33].)

Complete dry development was shown for bilevel resist schemes with an AZ-1450 photoresist base layer and an evaporated Ge top lithographic layer [22].

Because of the large depth of focus planarization is not a necessity for IPL pattern transfer.

Figure 11 shows a proposed process for pattern generation in metal layers using conformal organic (or inorganic [22,29]) resists. Here the ions penetrate the resist but do not degrade the properties of the underlying devices.

The use of high resolution inorganic resists and materials may become mandatory for sub-0.2-μm pattern replication because of the unstable high aspect ratio organic resist structures inherent to planarizing techniques (Fig. 10).

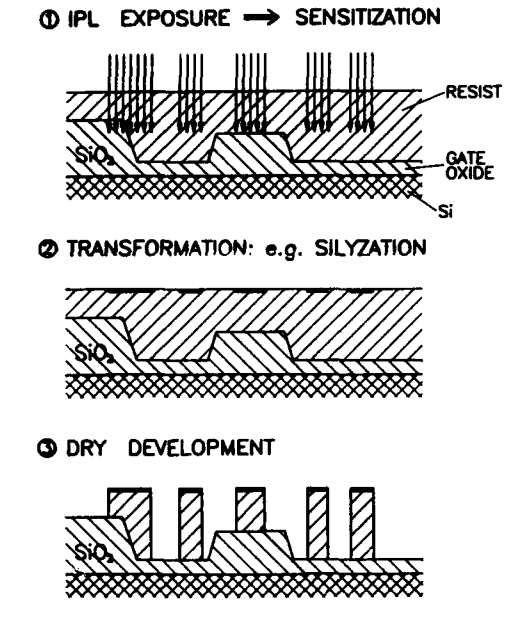


FIGURE 10. IPL pattern transfer into single layer organic resists.

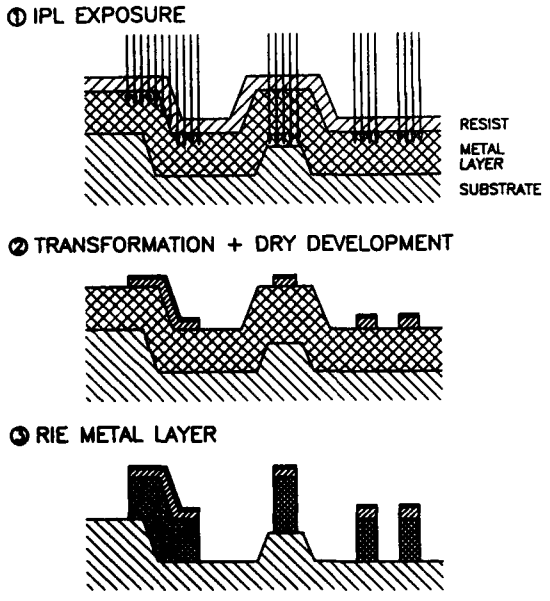


FIGURE 11. IPL pattern transfer into metal layer using conformal organic or inorganic resist layers.

4.5. IPL INTRAFIELD DISTORTION AND DIE-BY-DIE ALIGNMENT

First measurements of the intrafield distortion of the IPLM-01 ion-optical column show the possibility of changing the distortion from pincushion to barrel type [22].

This gives the possibility of an optimized set up of the ion-optical column by adjusting immersion and projective lens voltages and wafer Z-position.

In addition, with the help of ion-optical correction elements an electronic fine adjustment of intrafield distortion can be achieved [24,34]. Thus, a future "ion-optical" wafer stepper may be used in a mix & match mode together with optical, electron beam or x-ray lithographic tools.

With respect to die-by-die alignment a basic test has been performed studying signals of secondary electrons generated by ion beam probes sweeped electrostatically across wafer registration marks. The test was performed only in one direction using two channeltron detectors [24].

For chip alignment 8 channeltron detectors are necessary where a pair is used to generate signals for X, Y, rotation and scale. Because of the low energy of the ion probe induced secondary electrons [35] the wafer registration marks may not be covered by planarizing layers. (Organic resists may be selectively removed by excimer laser photoablation techniques [36].)

Resist removal can be avoided using combined ion projection and optical alignment techniques as schematically shown in Fig. 12.

With the help of channeltron detector signals the projected ion image can be pinned in X, Y, rotation and scale to a reference block situated between projective lens and wafer.

With an optical alignment system (e.g. using linear fresnel zone plates [37]) wafer registration marks can be detected.

The misalignment indicated by the optical system may be corrected by a mechanic wafer movement or - without moving mechanic parts - by an electronic correction of the projected ion image (in X, Y, rotation, and scale).

4.6. CONCLUSION

The results demonstrated so far with the IPLM-01 experimental setup lead to the conclusion that ion projection lithography can be developed to be an important high volume production technique for sub-0.5-μm devices.

Statistical data on resolution, pattern transfer characteristics, depth of field and their distributions over the image field have to be generated which holds also for intrafield distortion measurements.

Thus, in analogy to optical wafer and x-ray steppers, high speed (electrical) measurements techniques [38,39] must be implemented which can also be used for the evaluation of die-by-die alignment.

For many future devices (e.g. GaAs on Si) the large depth of focus of ion projection lithography will facilitate economic production techniques. IPL exposures may also be used for the development of in situ processing using ion beam modification techniques of materials [22,29].

The IPL mix & match possibilities may also lead to "chip composing" techniques using optical (respectively with higher process latitudes [40] synchrotron x-ray) 1:1 or optical reduction exposure techniques for $\geq$ 0.5 μm lithographic levels and demagnifying ion projection printing for levels requiring $\leq$ 0.5 μm resolution.

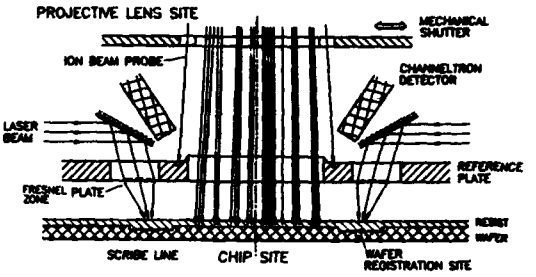


FIGURE 12. Principles of combined ion projection and optical alignment.

REFERENCES

[1] H. Ahmed, NATO Workshop on Emerging Technologies for In Situ Processing, Corsica, May 4-8, 1987, in print (North Holland).

[2] R.L. Kubena, R.J. Joyce, H.L. Garvin, F.P. Stratton, R.G.Brault, 31st Int.Symp. on Electron, Ion and Photon Beams, Woodland Hills, Ca, May 26-29, 1987, t.b.p. in J. Vac. Sci. Technol. (1988).

[3] M. Konishi, M. Takizawa, T. Tasumori, 31st Int.Symp. on Electron, Ion and Photon Beams, Woodland Hills, Ca, May 26-29, 1987, t.b.p. in J. Vac. Sci. Technol. (1988).

[4] L. Karapiperis, I. Adesida, C.A. Lee, E.D. Wolf, J. Vac. Sci. Technol. 19(4), 1259 (1981).

[5] H.I. Smith, Ext.Abstr. 18th (1986 Int.) Conf. on Solid State Devices and Materials, Tokyo, p.13 (The Japan Society of Applied Physics, 1986).

[6] W.L. Brown, NATO Advanced Study Inst., Heraklion, Crete, Sept. 16-22, 1985 (North Holland).

[7] B.M. Siegel, 31st Int.Symp. on Electron, Ion and Photon Beams, Woodland Hills, Ca, May 26-29, 1987, t.b.p. in J. Vac. Sci. Technol. (1988).

[8] K. Gamo, S. Namba, Mat. Res. Soc. Symp. Proc. Vol.45, p. 223 (Materials Research Society, 1985)

[9] J. Melngailis, 31st Int. Symp. on Electron, Ion and Photon Beams, Woodland Hills, Ca, May 26-29, 1987, t.b.p. in J. Vac. Sci. Technol. (1988).

[10] A. Wagner, Solid State Technol. 26(5), 97 (1983).

[11] B.M. Ward, D.C. Shaver, M.L. Ward, Proc. SPIE Vol. 537, p. 117 (1985)

[12] G. Stengl, H. Löschner, J.J. Muray, Solid State Technol. 29(2), 119 (1986).

[13] J. Orloff, P. Sudraud, Micro circuit Engineering 3, p. 161 (North Holland, 1985).

[14] G. Stengl, H. Löschner, W. Maurer, P. Wolf, J. Vac. Sci. Technol. B4(1), 194 (1986).

[15] B. Lischke, private communication.

[16] J.J. Muray, Semiconductor Int. 7(4), 130 (1983).

[17] J. Bartelt, Solid State Technol. 29(5), 215 (1986).

[18] G.M. Atkinson, J.L. Bartelt, A.R. Neureuther, N.W. Cheung, J. Vac. Sci. Technol. B5(1), 232 (1987).

[19] J.N. Randall, E.I. Bromley, N.P. Economou, J. Vac. Sci. Technol. B4(1), 10 (1986).

[20] H. Bohlen, U. Behringer, J. Keyser, P. Nehmiz, W. Zapka, W. Kulcke, Solid State Technol. 27(9), 210 (1984).

[21] J.N. Randall, R. Sivasankar, J. Vac. Sci. Technol. B5(1), 223 (1987).

[22] G. Stengl, H. Löschner, E. Hammel, E.D. Wolf, J.J. Muray, NATO Workshop on Emerging Technologies for In Situ Processing, Corsica, May 4-8, 1987, in print(North-Holland)

[23] G. Stengl. H. Löschner, J.J. Muray, Techn. Proc. SEMICON/West 1986, p. 42 (SEMI Inc., Mountain View, California, 1986).

[24] R.Fischer, E.Hammel, H.Löschner, G. Stengl, P. Wolf, H. Kraus, G. Stangl, Microelectronic Engineering 5, p. 193 (North-Holland, 1986).

[25] A. Heuberger, Microelectronic Engineering 5, p. 3 (North-Holland, 1986).

[26] A. Heuberger, L.M. Buchmann, L. Csepregi, K.P. Müller, this volume.

[27] F.O. Fong, C. Kinalidis, J.C. Wolfe, J.A. Oro, as [2].

[28] J.N. Randall, D. Bellavance, R. Sivasankar, Materials Research Society 1986 Fall Meeting, Boston, Dec. 1986, in print.

[29] G. Stengl, H. Löschner, P. Wolf, Nucl. Instr. Meth. in Phys. Res. B19/20, 987 (1987).

[30] S.W. Pang, T.M. Lyszczarz, C.L. Chen, J.P. Donnelly, J.N. Randall, J. Vac. Sci. Technol. B5(1), 215 (1987).

[31] F. Coopmans, B. Roland, R. Lombaerts, Microelectronic Engineering 5, p. 291 (North-Holland, 1986).

[32] W.L. Brown, Radiation Effects 98, 115 (1986).

[33] M.I.J. Beale, C. Broughton, A.J. Pidduck, V.G.I. Deshmuck, Nucl. Instr. Meth. in Phys. Res. B19/20, 995 (1987).

[34] G. Stengl, H. Löschner, J.J. Muray, Ext.Abstr. 18th (1986 Int.) Conf. on Solid State Devices and Materials, Tokyo, p.29 (The Japan Society of Applied Physics, 1986).

[35] A.J. Muray, J.J. Muray, Vacuum 35(10), 467 (1985).

[36] K.J. Polasko, D.J. Elliott, B.P. Piwczyk, as [2].

[37] B.S. Fay, W.T. Novak, Proc. SPIE Vol. 632, 146 (1986).

[38] K. Chivers, T. Hasan, Techn. Proc. SEMICON/Europa 86, p. 104 (SEMI Inc., Mountain View, California, 1986).

[39] B. Fay, J. Labrie, S. Bijawat, Microcircuit Engineering, 5, 587 (1986).

[40] H.I. Smith, 31st Int.Symp. on Electron, Ion and Photon Beams, Woodland Hills, Ca, May 26-29, 1987, t.b.p. in J. Vac. Sci. Technol. (1988).

Session C2.4

Power Devices II

Chairman: E. Rimini

Tuesday, September 15, 1987

EXPERIMENTS AND MODELING FOR U.H.F. POWER VERTICAL D.MOS TRANSISTOR

G. TARDIVO * - P. ROSSEL ** - A. SENES * - G. CAZAUBON *
M. BELABADIA ** - R. MAIMOUNI **

* THOMSON SEMICONDUCTORS
Rue P. et M. Curie - BP 0155
37 001 TOURS Cedex - FRANCE

** LAAS - CNRS
7, Avenue du Colonel Roche
31 077 TOULOUSE Cedex

ABSTRACT

After a brief description of the basic structure and fabrication process of V.D.MOSFET's for power high frequency applications, a new non-linear dynamic model is proposed. The implantation of this model on the ASTEC III simulation program, enables its use for comparison between measured and computed characteristics under DC and small - signal conditions and to set up a new simulation procedure for designing and simulating radio-frequency power amplifiers. At last, the model is used to give some design rules for U.H.F. power V.D. MOSFET's.

1. INTRODUCTION

Since it came out on the radio-frequency power semiconductor market the MOSFET has played an increasing role, because of its numerous advantages like good power gain, high input impedance, and very good thermal stability [1-6]. Among MOSFETs, the vertical double-diffused MOS (V.D.MOS) technology is now chosen by almost all manufacturers. Moreover, beside the technological developments, it appears necessary to set up modelization and simulation tools both to define design methodologies that apply to radio-frequency power amplifiers using these transistors and to obtain design rules for future V.D.MOSFETs operating in the U.H.F. range.

2. BASIC STRUCTURE AND FABRICATION PROCESS

The V.D.MOS transistors for radio-frequency applications are vertical devices, that is to say the drain contact is located at the bottom of the chip, whereas the source and gate contacts define an interdigitated pattern on the top of the structure. Figure 1 shows a cross-sectionnal view of this structure.

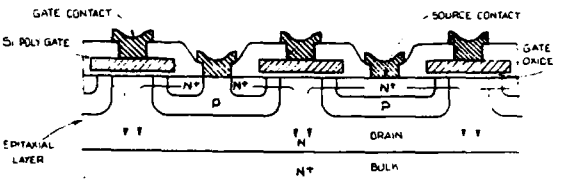


FIGURE 1

Cross - sectionnal view of a V.D. MOSFET.

The starting material is an epitaxial silicon layer whose thickness and resistivity are carefully chosen in order to increase the device voltage handling capability, without damaging the on-state characteristics. Figure 2, shows the main steps of the fabrication process. This process begins with a gate oxidation, followed by a polysilicon layer deposition, which will be etched in order to create the gate pattern. A p-type source region will be diffused, always using the gate pattern as part of the mask. This double-diffusion process is self-aligned on the gate, so as to enable an accurate control of the channel length (lower than 2 microns). The polysilicon gate is then coated with an aluminium metallization and the source metallization is realized. Finally, a passivation film is formed. Chips are packaged in a 4L FL package by using Au/Si solder.

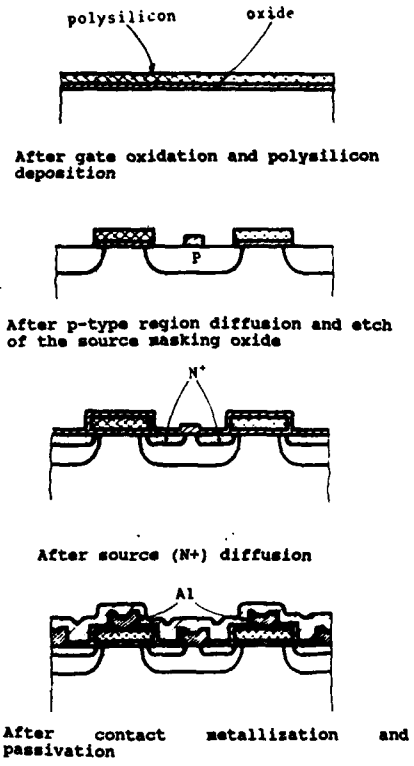


FIGURE II

Main technological steps of the V.D. MOSFET fabrication process.

3. MODELING OF THE V.H.F. V.D.MOS TRANSISTOR

An accurate high frequency model for the V.D.MOS transistor must take into account all the specific characteristics of this device and more precidely :

i/ the short channel effects on the scattering velocity,
ii/ the lumped configuration of the inverted channel and the accumulated region under the gate
iii/ the current spreading effects in the drain epilayer,
iv/ the influence of the drain diode and overlap capacitances as well as parasitics elements due to encapsulation.

3.1. Modeling of the channel region

In contrast with the "small signal" models proposed up to now, the model we propose for the channel region is based on a mathematical resolution of the general semiconductor equations, taking into account the saturation of the scattering velocity in both lateral and transverse directions. Detailed calculations are described in references [7, 8]. The active channel region is treated as a non-uniform transmission line. The first differential dynamic equation is obtained by linearizing the expression of the current and the second expresses the conservation of the current flow taking into account the current losses through the gate. After development of the equations'solutions, it appears that the channel region can be modelized by a discrete element equivalent circuit, with capacitance and resistors whose values only depend on the geometrical and physical parameters and on the instantaneous voltages. Figure 3 shows the dynamic equivalent scheme of the channel region.

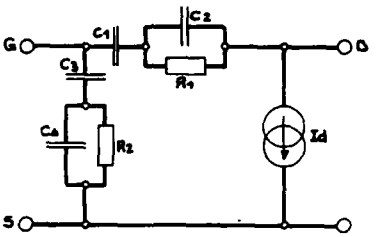


FIGURE III

Equivalent scheme for the channel region of the V.D. MOSFET.

3.2. Model of the V.D. MOS transistor

The second main region of the V.D.MOS transistor is the epilayer region between the two channels under the gate. We have shown, using bidimensionnal simulations, that this region could be either accumulated or depleted A dynamic equivalent lumped configuration is proposed and the involved equations are solved (detailed calculations in réf. [9]) to obtain a simple equivalent circuit.

The current spreading effect in the drain epilayer is taken into account as a variable resistor given by an analytical formula [10]. All the capacitances (diffused jonctions, overlap, ...) are modelized by simple classical formulas. Lastly parasitics elements such as metallization resistors, bonding wires inductors and inter-metal capacitances are added to the model. Figure 4 shows the complete equivalent scheme.

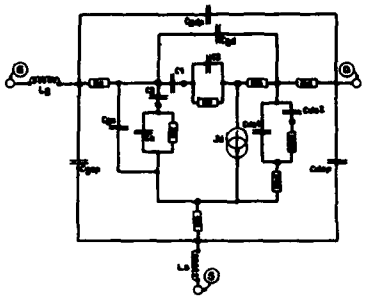


FIGURE IV

Complete Dynamic Equivalent Scheme for the V.D. MOSFET.

4. DC AND "SMALL SIGNAL" SIMULATIONS

The model is implanted on the non-linear circuit simulator ASTEC III. It allows the determination of the following device characteristics : DC drain current versus both gate and drain voltages, scattering parameters, power gain, LINVILL stability factor and switching characteristics.

The computed parameters were compared to their measured values [9]. Figure 5 shows that comparison for the S21 scattering parameter. The difference between measurement and simulation proves to be better than 30 %. This precision is typical for this kind of model and due to measurement imperfections, can be regarded as a good validation for our model.

5. DESIGN AND SIMULATION OF RF POWER AMPLIFIERS

One of the main application of the non-linear model of the V.D.MOS transistor is the setting-up of a new procedure for designing and simulating radio-frequency power amplifiers, based on the determination of all the large signal parameters (i.e output power compression point, input and output impedances, large signal power gain, power efficiency).

This new procedure is based on the numerical determination both of the input and output power deduced from the computation of the instantaneous current and voltage values, in the time domain. First, the device is unconditionnally stabilized for all gate and drain bias conditions using series, or parallel or feedback network [9] ; second, the large signal input and output impedances and the maximum available output power from the device are obtained [11]. It is so possible to obtain the device performances under A, B or C working classes in determining output power versus input power characteristics and the large signal power gain.

Figure 6 shows a comparison between experiments and simulation for a 45 watts devices realised using the previous described technology. Several comparisons of this kind were realized and a very good agreement was noticed in each case, between measured and calculated curves (difference lower than 1.5 db).

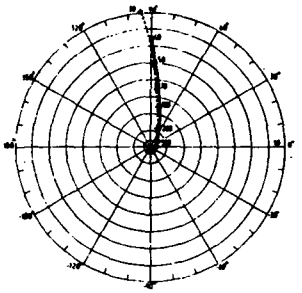


FIGURE V

Comparison between the experimental and computed curves for the S 21 scattering parameter (45 watts - MOSFET - 2/200 MHz/____ experiments --- simulation).

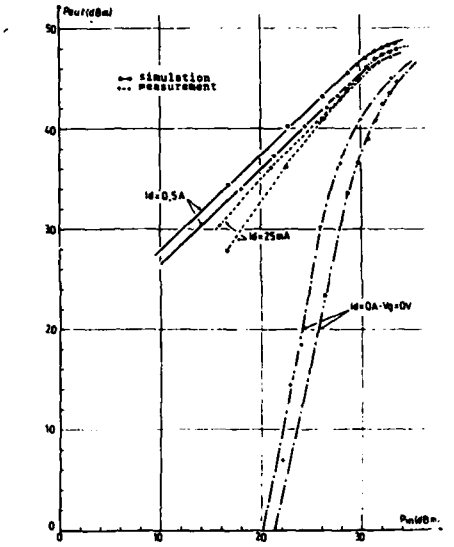


FIGURE VI

Comparison between the experimental and computed RF large signal (A, B and C classes) characteristics. Output power versus Input power for a 45 watts V.D. MOSFET - Vds = 28 V - F = 150 MHz.

6. ELEMENTS FOR DESIGNING U.H.F. POWER V.D. MOSFET's

These model and procedures could be easily used for the design of new devices and performance predictions. As an example, we have studied how a V.D.MOSFET should be designed to have a good efficiency as an U.H.F. power amplifier. We have shown that, by reducing the size of the source cells, minimizing the source inductance and choosing a partly overlapping gate configuration, very good U.H.F. performances are obtainable [9].

Figure 7 shows the evolution of the output power versus input power characteristics as a function of the rate of overlapping gate removed. From a simulation point of view, we prooved that U.H.F. power V.D., MOSFET's can be designed with performances as good as a power gain of 17 db at 500 MHz and 10.5 db at 1 GHz for an output power of 80 watts.

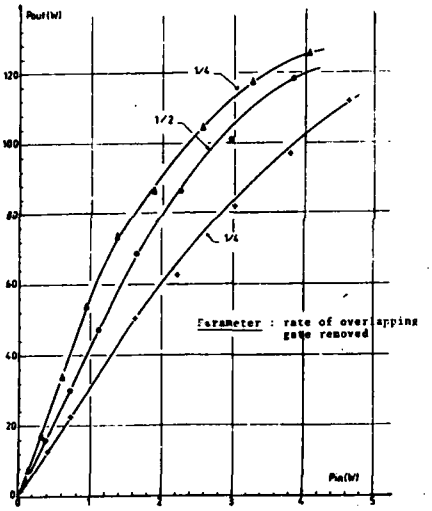


FIGURE VII

Computed large signal characteristics for V.D.MOSFET's with partly overlaping gate - Vds = 28 V - Output power : 80 W F = 500 MHz - Id = 0,5 A.

7. CONCLUSION

As a conclusion, we have proposed beside a new technology development, a new non-linear model for the V.D.MOS transistor. This model proves to be well adapted both to simulate the device characteristics and to design radio-frequency power amplifier using these transistors. Moreover, it enable the design of new generation V.D.MOSFET's operating in the U.H.F. range, whose characteristics are predictible before any technological realization.

REFERENCES

[1] H.J. SIGG, G.D. VENDELIN, T.P. CAUGE, J. KOCSIS, "D.MOS Transistor for Microwave Applications". I.E.E.E. Trans. on E.D., vol. ED - 19, n° 1, January 1972.

[2] E. FONG, D.C. PITZER, R.J. ZEMAN, "Power D.MOS for High - Frequency and Switching Applications", I.E.E.E. Trans. on E.D., vol. ED - 27, n° 2, February 1980.

[3] K. SEDIGH and Al. "D.MOS FET produces 100 W at 400 MHz". Microwaves and RF, november 1985.

[4] H. ESAKI, O. ISHIKAWA, "A 900 MHz 100 W VD - MOSFET with silicide gate self-aligned channel", I.E.D.M. 1984.

[5] O. ISHIKAWA, H. YAMADA, H. ESAKI, "A 2.45 GHz Power L.D. MOSFET with reduced source inductance by V-groove connections", I.E.D.M. 1985.

[6] J. MENA, "High Frequency Performance of VDMOS Transistors", M.A. Sc Thesis, TORONTO, 1981.

[7] L. BELABAS, "Propriétés en hautes fréquences du transistor métal - oxyde - semiconducteur à canal vertical (VMOS)". Thèse de docteur ingénieur, TOULOUSE, 1983.

[8] S. LATRECHE, G. TARDIVO, M. BELABADIA, P. ROSSEL "Modèle mathématique du transistor V.MOS en régimes linéaires et non linéaires. Logiciels pour les simulations en commutation et en amplification haute-fréquence de puissance". Rapport LAAS n° 85.112, mai 1985.

[9] G. TARDIVO "Le transistor D.MOS vertical en amplification haute fréquence de puissance". Thèse de 3ème cycle - TOULOUSE - 1987

[10] J.L. SANCHEZ, "Propriétés à l'état passant des transistors D.MOS de puissance coplanaires ou verticaux". Thèse de Docteur Ingénieur, TOULOUSE, 1984.

[11] S. LATRECHE, G. TARDIVO, M. BELABADIA, P. ROSSEL "Amplification de puissance en régime non linéaire et en hautes fréquences à transistor MOS". Revue de physique appliquée, mai 1987.

References [7, 8, 9, 10] are available on request.

VERIGRID-FCTh SWITCHING 10A AT 1000V

H. Gruening and J. Voboril

Research Center, CRBS.L
BBC Brown Boveri AG
CH-5405 Baden Switzerland

New field controlled thyristors (FCTh) have been realized by a recessed gate technique (VERIGRID), which exhibit a higher aspect ratio of the control fingers than devices produced so far. Thus a static blocking gain in excess of 500 could be achieved, and more than 10A (125A/cm^2) were switched off snubberless at 1000V and inductive load. Furthermore the doping of the finger sidewalls turned out to be very important for the on state: with high doping a hole bypass is created, and long channel JFET saturation is observed instead of a low resistance like that of a pin diode.

1. Introduction

Since the Field-Controlled Thyristor (FCTh) has been proposed by Nishizawa e.a. [1], different approaches have been made to realize such a device [2,3,4]. Nevertheless the ultimate limits of the concept have not been shown so far, and, instead of this, nearly all the efforts have been put on the IGBT, since it seemed to be the functional integration of the FCTh-MOSFET cascode [4].

Now IGBTs have been developed by Nakagawa e.a. [5,6], and an effective hole bypass had to be implemented in the cathode region to prevent latchup. Computer simulations of this device [5] show that in the on state carriers are injected from the anode only. Thus an FCTh could exhibit lower on resistance and switching losses for two reasons: First there is no MOSFET connected in series with the bipolar high voltage device, and secondly, carrier injection from both sides, anode and cathode, should be possible within an FCTh.

Our recessed gate technology (VERIGRID) recently developed for a highly interdigitated GTO [7] turned out to be the key factor for the realisation of FCThs. A new latching mechanism, the dynamic latching, has been found [8], and, from experimental data, we proposed the aspect ratio of the cathode fingers to control this mechanism. Recent results on this statement as well as data giving a deeper understanding of the physics of the on state are presented in this paper.

2. Design and Technology

The structure of a test device is shown in Fig.1. It is made from FZ-Si with anode, buffer layer and termination as usual. Cathode fingers are etched by reactive ion etching (RIE) after diffusion of the n^+ regions, and the surfaces of the grooves are doped with boron. Metallic contact to the cathode stripes and the bottom regions of the gate grooves is realized by evaporated Al, and passivation of the termination moat as well as isolation between gate and cathode is established by photopolyimide. During a last processing step, a solderable metal layer is sputtered on top of the cathode region and contacted by a copper plate. No lifetime control has been applied at the elements presented within this paper.

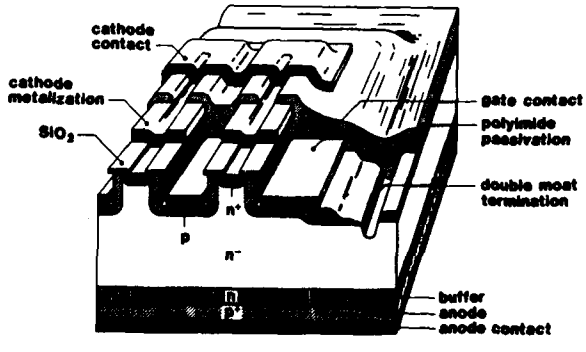


FIGURE 1

Schematical structure of a test FCTh. The 102 cathode fingers (25 μm x 1.3mm x 30μm) form an active area of 8mm^2.



FIGURE 2

SEM picture of the FCTh cathode region.



FIGURE 3

On state characteristics of FCTh with low gate doping.

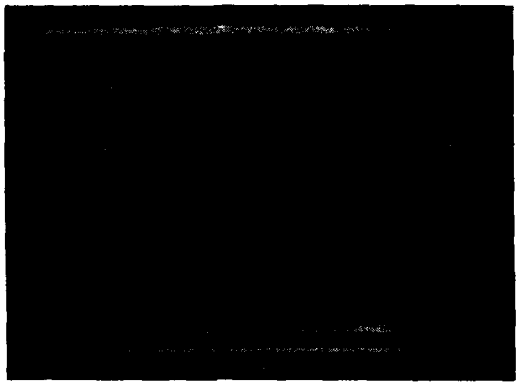


FIGURE 4

On state characteristics of FCTh with high gate doping:
a) zero gate current,
b) 3mA positive gate current.

3. Results and Discussion

3.1 On State

With the aspect ratio near one realized today two different characteristics with no gate current applied are observed: devices with low gate doping behave like diodes (Fig.3), while elements with higher doping exhibit pentode like saturation (Fig.4a). Then a small gate current is necessary to obtain a turn on with low on resistance (Fig.4b).

Since there are no detailed simulations of the FCTh available so far, to obtain information about this effect, the gate current has been measured as a function of gate cathode voltage and anode current. Thus it was found that only a very small decrease in gate potential from the on state is necessary to extract all the holes drifting through the bulk. Then conduction by electrons only is left in the cathode fingers, and long channel JFET operation results, as shown in Fig.5.

But a decrease in gate potential also may be due to an internal bypass: high gate doping causes high conductivity along the finger wall. Then the holes run from the bottom of the gate along the wall, and across the p-n junction at the top towards the cathode (Fig.5). Of course, internal and external gate currents create a voltage drop at this wall resistor, and at a certain net value of both enough holes will have to go through the channel again. Thus the thyristor like characteristic of Fig.4b not really results from a four layer structure, but from a switchover from unipolar to bipolar conduction in the FCTh's fingers.

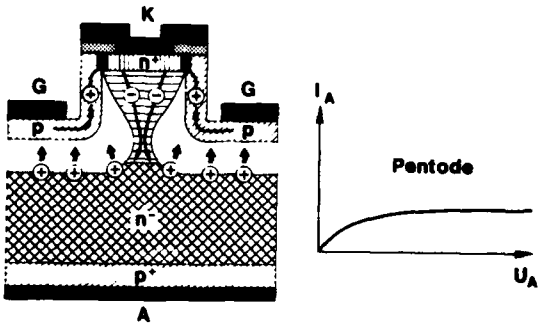


FIGURE 5

Schematic of hole bypass created by high gate doping. This causes unipolar JFET operation in the cathode fingers.

3.2 Static Blocking

As expected, a drastic increase in blocking gain has been observed with respect to results published recently [8]. Now more than 1500V can be blocked by -4V at the gate only (Fig.6). With anode currents in the μA range, the blocking characteristics still look triode like, but in the mA range they saturate very similar to that shown in Fig.4a.

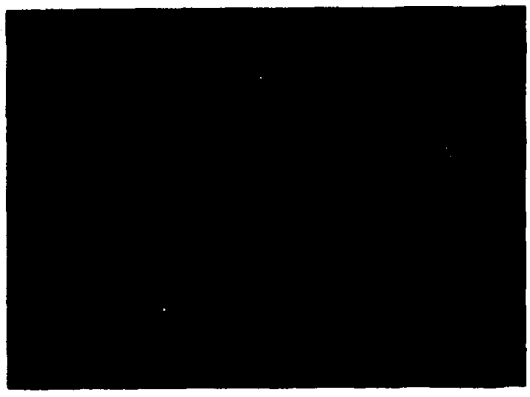
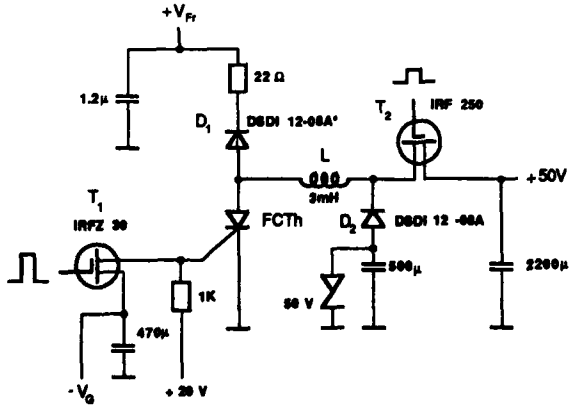


FIGURE 6

Blocking characteristics of typical FCTh:
V_{GK}= 0V,-1V,-2V,-3V,-4V,-5V.



* Diode selected for $V_{DRM} \geq 1500$ V

FIGURE 7

Non-destruct RBSOA tester.

3.3 Switching On and Off

Similar to our nondestructive test circuit with resistive load [8] we have developed a nondestructive test for inductive load (Fig.7). One test cycle then consists of four parts:

1. FCTh and T_2 are turned on, and the current through the inductor L is increasing until the test current is reached.
2. The FCTh is turned off. Now its anode voltage rises until D_1 starts conducting.
3. When testing the SOA, the FCTh is switched on again about 0.5μs later. Therefore no destruction can occur in case of insufficient turn off.
4. T_1 is switched off again, and the energy still stored in the inductor L now is transferred to a zener by D_2.

A switching result is shown in Fig.8. An anode current of 10A is switched off at V_{Fr}= 1000V, and the anode voltage increases up to 1200V because of the 22Ω resistor connected in series with D_1. As expected from the carrier lifetime (2μs), quite a long tail is observed during turn off. On the other hand turn on is accomplished very fast, although there is only a 1kΩ pull up resistor at the gate of the FCTh. In fact, holes injected from the anode during the onset of conduction first decharge the gate. Thereby an internal feedback causing rapid turn on of the FCTh is generated in contrast to the miller capacitance, which is slowing down the switching of MOSFETs or IGBTs.

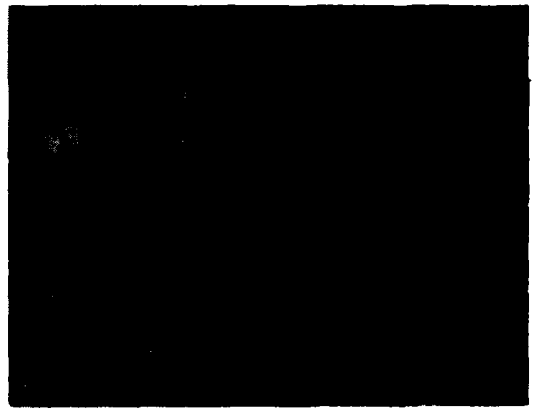


FIGURE 8

Switching wafeform obtained at
V_{Fr}= 1000V and I_A= 10A.
V_A: 500V/Div, I_A: 4A/Div,
Time: 2μs/Div.

For a detailed study of the turn off process Fig.9 is presented. Although parasitics could not be avoided totally, four steps of the process clearly are distinguishable:

1. Rise in gate current with constant di/dt, accompanied by cathode current reversal,
2. abrupt increase in negative gate voltage,
3. stop of cathode current flow and
4. increase in anode voltage with all the anode tail current flowing along the gate.

Therefore we conclude that during FCTh turn off first the cathode fingers are depleted, completly interrupting the injection of electrons from the cathode. Then the depletion layer homogeniously starts growing from the gate towards the anode as within an IGBT, for instance. Thus current crowding should be minimized.

4. Conclusion

We presented an FCTh which at the first time is snubberless switching 1000V at 125A/cm^2 inductive load. No avalanche is observed so far indicating a very homogenious current spread during turn off. Experimentally a detailed design of the FCTh control structure is obtained, showing the necessity of the VERIGRID technology to realize FCThs. Thus, using our advanced FCTh cascade [8], a new, very high gain snubberless power switch is obtained exhibiting the rudgeness of bipolar devices and, potentially, an better switching speed for a given on state loss than IGBTs.

5. Acknowledgement

We like to thank Ch. Abbas, B. Broich and P. Roggwiller for their support and fruitful discussions, A. Blatter for his contributions to RIE and photopolyimide and Ch. Laabs, St. Mair and H. Keser for their technical assistance.

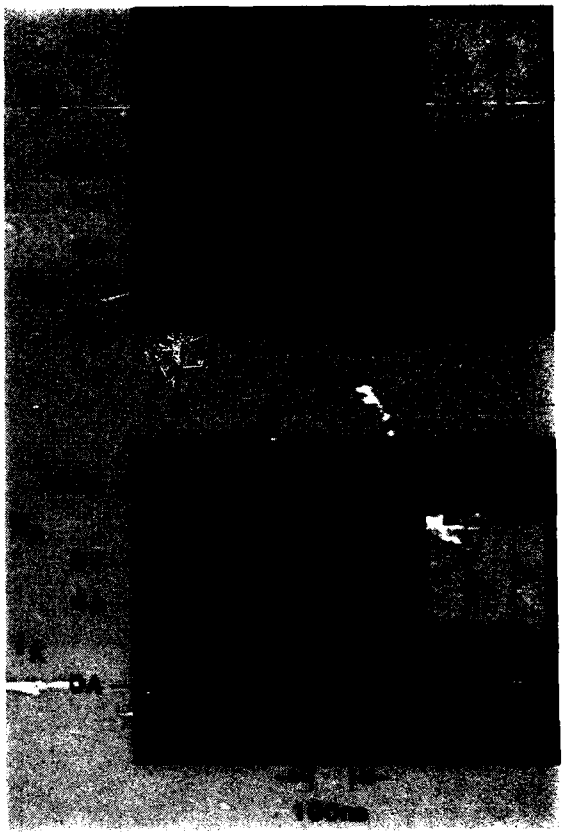


FIGURE 9

Typical turn off process of FCTh with resistive load.

6. References

[1] J.I. Nishizawa, T. Terasaki and J. Shibata, IEEE Trans. Electron Devices, Vol. ED-22, pp. 185-97, 1975.

[2] B.W. Wessels and B.J. Baliga, IEEE Trans. Electron Devices, Vol. ED-25, pp. 1261-65, 1978.

[3] B.J. Baliga, IEEE Trans. Electron Devices, Vol. ED-27, pp.1262-68, 1980.

[4] B.J. Baliga, Solid-State Electronics, Vol. 25, pp. 345-353, 1982.

[5] A. Nakagawa, Y. Yamagushi, K. Watanabe, H. Ohashi and M. Kurata, IEDM 1985, pp.150-153.

[6] A. Nakagawa, K. Watanabe, Y. Yamagushi, H. Ohashi and K. Furukawa, IEDM 1986, pp. 122-125.

[7] P. Roggwiller, J. Gobrecht, J. Voboril and B. Broich, IEDM 1984, pp. 439-442.

[8] H. Gruening, J. Voboril, J. Gobrecht, P. Roggwiller, C.C. Abbas and B. Broich, IEDM 1986, pp. 110-113.

ELECTRON ENERGY EFFECTS IN B-IRRADIATED POWER SEMICONDUCTOR DEVICES *

P.G. Fuochi*, A. Martelli*, E. Gombia**,
C. Malfatto***, B. Passerini***, M. Zambelli***

*) Istituto FRAE - CNR, Via de' Castagnoli 1, 40126 Bologna, Italy
**) Istituto MASPEC - CNR, Via Chiavari 18/A, 43100 Parma, Italy
***) ANSALDO, Semiconductor Dept., Via Lorenzi 8, 16152 Genova, Italy

Fast recovery power p-i-n diodes, obtained from <1:1> NTD float-zone silicon, have been irradiated at room temperature with monoenergetic electrons at 6.6, 9.2 and 11.7 MeV. The absorbed doses ranged from 2 to 16 kGy. The complex deep level spectra, revealed by DLTS measurements, have been related to the energy and dose values. The tradeoff between the static and the dynamic characteristics of the devices vs. the irradiation conditions have been investigated.

1. INTRODUCTION

Studies carried out in the past on radiation effects in silicon have found practical application in recent years, when a great interest has developed in the use of gamma irradiation and particularly high-energy electrons for recombination lifetime control in silicon devices [1,2]. The feasibility of controlling the electrical characteristics of power semiconductor devices by irradiation with high-energy electrons has been widely investigated [3]. However, little work has been done on the efficiency of electron irradiation when using electron energies in the range 5 - 12 MeV. The aim of this work is to investigate the influence of 6.6, 9.2 and 11.7 MeV electron beam energies at different doses both on the static and the dynamic characteristics of power p-i-n diodes. Also reported are measurements of the dependence of the room temperature production of the energy levels inside the silicon forbidden gap vs. the electron energy.

2. EXPERIMENTAL DETAILS

2.1 Device Fabrication

Fig. 1 shows the cross-sectional structure of a power p-i-n diode. The device is essentially a full diffused silicon wafer alloyed onto a molybdenum backplate which grants the necessary mechanical support. The silicon diodes used in this study were fabricated upon 120 ohm.cm neutron transmutation doped, float-zone silicon (FZ-NTD). The silicon wafers, 31 mm in diameter, were processed through standard steps, normally used in power semiconductor processing and described previously [4], in order to obtain the p-i-n structure. The p^+ layer is gallium doped with a surface concentration of 10^{18} cm^{-3} and the n^+ layer is phosphorus doped to 10^{19} cm^{-3}. After diffusion of gallium and phosphorus the wafers were alloyed onto molybdenum backplates. Metallization, contour bevelling and junction passivation gave finished devices ready for irradiation.

* This work has been supported by the Italian National Research Council (CNR) within the Finalized Poject "Materiali e Dispositivi per l'Elettronica allo Stato Solido".

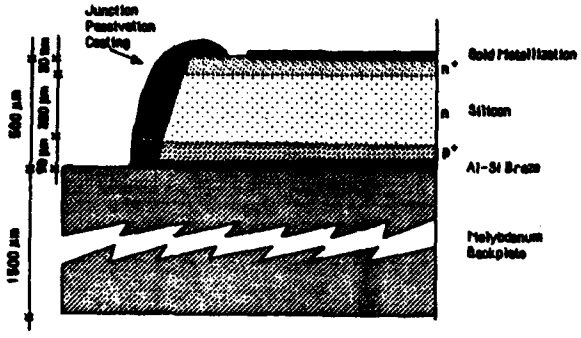


FIGURE 1

Cross-sectional structure of a power p-i-n silicon rectifier.

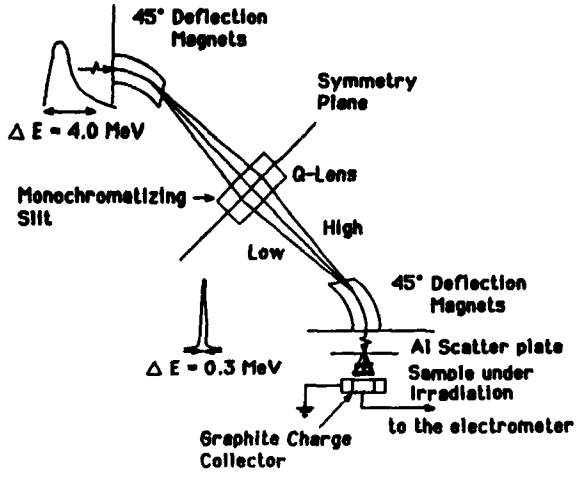


FIGURE 2

Schematic view of the experimental arrangement used for electron irradiation.

2.2 Electron-Irradiation

The p-i-n diodes were irradiated at room temperature using pulses of monoenergetic electrons from an L-band 12 MeV linear accelerator. The accelerator was used in repetitive mode with pulses of 50 ns duration and repetition rate of 600 p.p.s. Different electron energies were obtained by deflecting the beam by means of two 45 degrees magnets and by monochromatizing it by a slit placed in the maximum dispersion zone (Fig. 2). The resulting energies were 6.6, 9.2 and 11.7 MeV with a beam energy spread of +/- 0.15 MeV. The devices were irradiated with doses between $3x10^{12}$ and $3x10^{13}$ $e^-.cm^{-2}$. Fluxes of several 10^{11} $e^-.cm^{-2}.s^{-1}$ at the sample position were used and this guaranteed that the sample temperature did not exceed 50 °C during irradiation. Further details on the irradiation process and dosimetry using the Fricke chemical dosimeter have been already reported [4].

2.3 DLTS Measurements

DLTS measurements were carried out using a high sensitivity lock-in type spectrometer. A normal rate window of 952 s^{-1} was selected. It can be varied over more than three decades for studying the temperature dependence of the thermal emission rate. The DLTS measurements were performed by using a reverse bias of -10 V and a typical exciting pulse width of 100 μs. The trap density N_t was obtained by using a modified Zotha-Watanabe expression [5].

2.4 Electrical Characteristics

The effects of electron irradiation on the electrical characteristics of the device have been assessed by measuring the forward voltage drop V_F, the reverse recovery time t_{RR} and the recovered charge Q_{RR}.

3. RESULTS AND DISCUSSION

3.1 DLTS Spectra vs. Irradiation

DLTS spectra of unirradiated p-i-n samples (Fig. 3a) show only a broad band of electron traps centered at about 210 K. From the capacitance

(E_1, E_2, E_3 in Fig. 3) are clearly distinguishable. The E_1 level is the well characterized A-center (O-V pair) located at 0.17 eV from the conduction band. The E_2 (E_c-0.23 eV) and E_3 (E_c-0.42 eV) have been identified as the double negative $[V-V]^{--}$ and single negative $[V-V]^{-}$ charge state of the divacancy respectively [2,6]. From measurements performed on several samples we found that the concentration of these defects depend both on the energy and absorbed dose. In fact, at the same electron beam energy, the deep level density increases with the absorbed dose from 1.9 to 15.8 kGy while for a fixed dose an increasing of the trap concentration from 6.6 and 9.2 MeV was found. At 11.7 MeV no significative variation in the defect introduction rate was observed with respect to 9.2 MeV. Moreover no appreciable change in the relative concentration of the A-center and the divacancy has been found on samples irradiated with electron beam energies ranging from 6.6 to 11.7 MeV.

The ratios between the densities of these two centers, obtained on different samples irradiated at 6.6, 9.2 and 11.7 MeV are very close to the values reported by Brotherton and Bradley [7] for broad electron beam having 12 MeV maximum energy.

3.2 Electrical Characteristics vs. Irradiation Conditions

The forward voltage drop as a function of dose for the three different energies used is shown in Fig. 4. As reported elsewhere [8], the forward characteristics are controlled by the high level lifetime: the lower this value, the higher V_F. The massive stopping power (and hence the radiation damage) of the device decreases as the electron energy is increased [9]. Therefore the charge carrier lifetime is expected to be a decreasing function of the electron energy. This accounts for the relationship between the electron energy and V_F. The same considerations explain the behaviour of the dynamic characteristics t_{rr} and Q_{rr}, whose

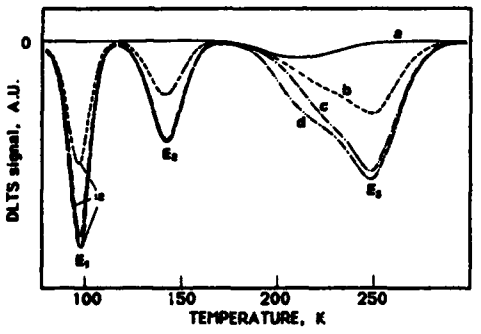


FIGURE 3

DLTS spectra from float-zone, NTD silicon samples before (a) and after electron irradiation electron energies: 6.6 MeV (b), 9.2 MeV (c) and 11.7 MeV (d); electron dose = 3.4×10^{12} e^- cm^{-2} (1.9 kGy in Fricke dosimeter).

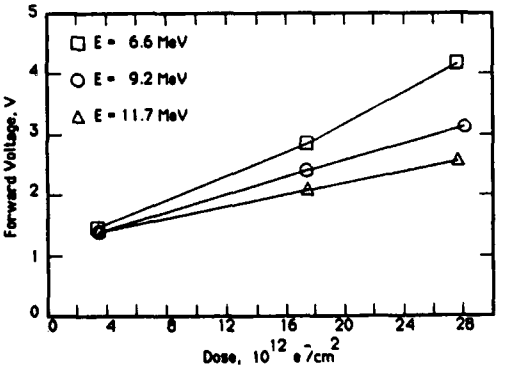


FIGURE 4

Forward voltage drop V_F vs. electron fluence, at different electron energies. Measuring conditions: I_F= 1200 A, T_j= 25°C.

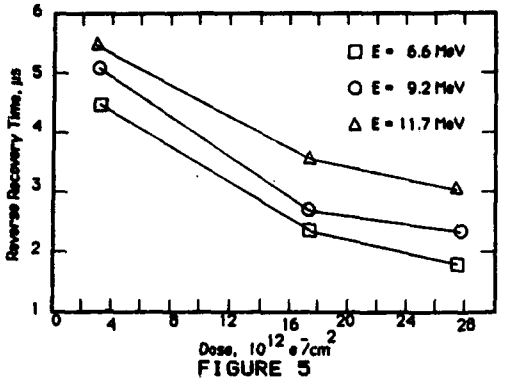


FIGURE 5

Reverse recovery time t_{RR} vs. the electron fluence , at different energies. Measuring conditions: I_F= 350 A, dI/dt= -80 A μsec^{-1}, V_R= 600 A, C_S= 0.25 μF, R_S= 10 ohm, T_j= 150 °C.

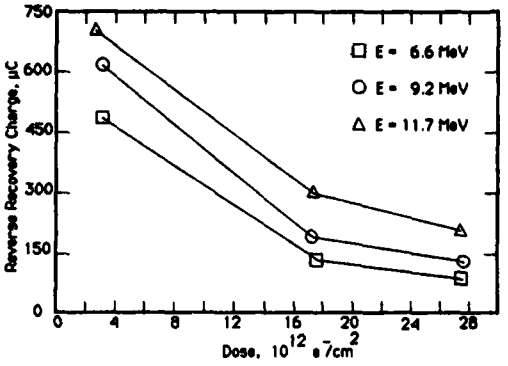


FIGURE 6

Reverse recovered charge Q_R vs. electron fluence, at different energies. Measuring conditions: I_F= 350 A, dI/dt= -80 A μsec^{-1}, V_R= 600 A, C_S= 0.25 μF, R_S= 10 ohm, T_j= 150 °C.

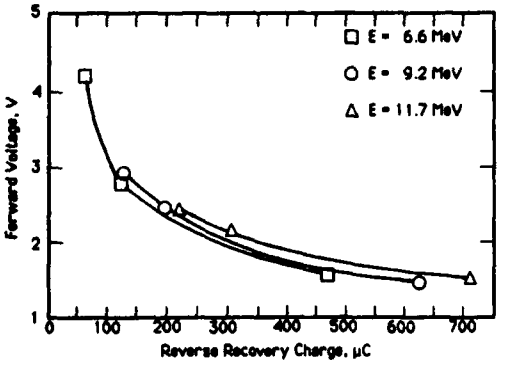


FIGURE 7

Trade-off curves between forward voltage drop (V_F) and recovered charge (Q_R) at different energies.

variation with energy and dose are reported in Figg. 5 and 6, respectively. Fig. 7 shows the V_F - Q_{RR} trade-off curves at the various electron energies used. The differences among these curves are not very pronounced. Anyway they seem to suggest the 6.6 MeV energy are more favourable compared to the 9.2 and 11.7 MeV.

ACKNOWLEDGEMENTS

The authors are in debt with Dr. R. Mosca for the DLTS measurements.

REFERENCES

[1] Carlson, R.P., Sun, Y.S., and Assalit, H.B., IEEE Trans. Electron Dev., vol. ED-24 (1977), 1103.

[2] Evwaraye, A.O., and Baliga, B.J., J. Electrochem. Soc., vol. 124, (1977), 913.[4] IEC Rec., Publ. 147-0 (1966), 37.

[3] Fuchs, H., Grube, R., Knopp, J. and Tursky, Electron Irradiation of Semiconductor Devices, (NTIS, U.S. Dept. of Commerce, Jan. 1983) and references therein.

[4] Fuochi, P.G., DiMarco, P.G., Bisio, G.M., DiZitti, E., Passerini, B., Tenconi, S., and Zambelli, M., Alta Frequenza, vol. 55 (1986), 47.

[5] Ghezzi, C., Gombia, E., and Vanzetti, L., Mat. Science Forum, 10-12 (1986), 1213.

[6] Evwaraye, A.O., and Sun, E., J. Appl. Phys., vol. 47 (1976), 3776.

[7] Brotherton, S.D., and Bradley, J., J. Appl. Phys., vol. 53, no. 8 (1982), 5720.

[8] Baliga, B.J., and Krishna, S., General Electric Techn. Inf. Series, Rep. no. 75CRD251 (1975).

[9] Berger, M.J., and Seltzer, S.M., Nat. Acad. of Sci. - Nat. Res. Counc., Publ. 1133, Nucl. Sci. Series Rep. no. 39

Helium Implantation for Lifetime Control in Silicon Power Devices

Wolfgang Wondrak and Alfred Boos

AEG Forschungsinstitut Frankfurt, Goldsteinstr.235, 6000 Frankfurt 71

The concentration profile of defect centres in silicon induced by α -particle irradiation, is investigated using deep level transient spectroscopy and spreading resistance measurements. As in the case of proton irradiation, a buried recombination centre doped layer is created in the penetration depth of the particles. In contrast to proton irradiations, after annealing of α -particle irradiated samples no shallow donor doped layers are observed, which can severely affect the breakdown voltage of power devices.

1. INTRODUCTION

High energy (> 1MeV) implantation of protons in silicon results in narrow radiadion damaged layers buried in depths > 10 µm. This has been discussed as a new technology for localized lifetime reduction in power devices, such as diodes and gate-turn-off thyristors[1], insulated gate transistors[2], and static induction thyristors[3]. Additionally to the formation of recombination centres, proton implantation and subsequent annealing introduces shallow hydrogen donors which affect the original doping profiles of the devices. This may be advantageous in some cases, but for power diodes for example, a recombination layer with enhanced n-type conductivity located near the pn-junction will reduce the blocking voltage (see fig. 1). In order to overcome this problem, α -particle irradiation has been proposed[4].

In this paper, the defect distributions in proton- and in α-particle irradiated silicon are compared. We show, that for localized lifetime control in power diodes α-particle irradiation is superior to proton irradiation.

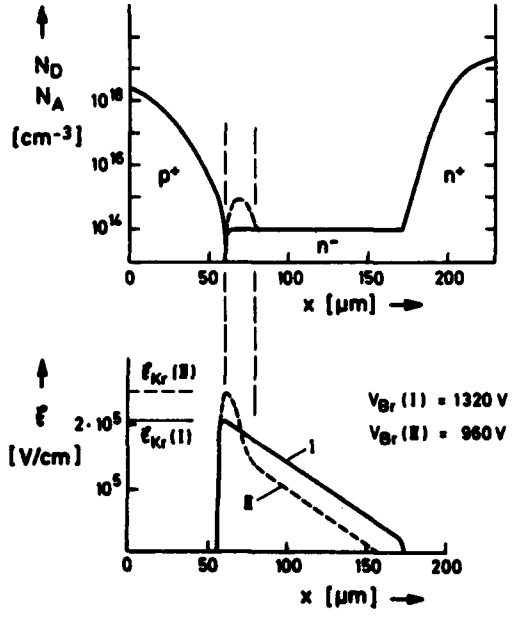


FIGURE 1

Influence of a buried n-doped layer on the breakdown voltage of a p^+-n^--n^+ power diode (calculated)

2. RESISTIVITY PROFILES

Integral information on the defect concentration profiles was obtained by spreading resistance measurements. Wacker-float-zone silicon doped with 10^{14} cm^{-3} phorphorous was irradiated with protons (2-5 MeV) and α-particles (3-11 MeV). After proton irradiations (fig. 2) and after α-particle irradiations (fig. 3), the resistivity increases. The high resistivity peak is always followed by a layer with slightly increased conductivity.

These investigations have been supported financially by the "Ministerium für Forschung und Technologie", FRG

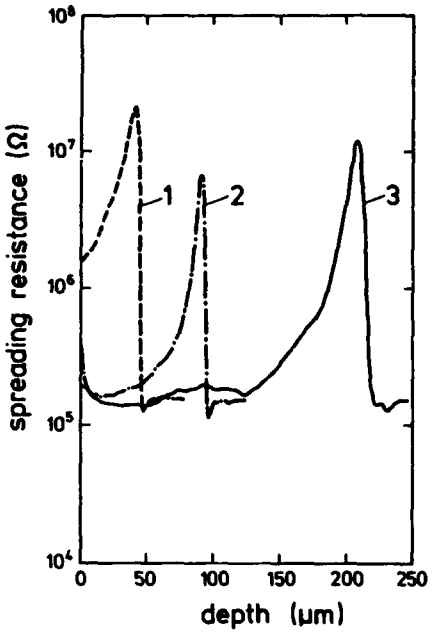


FIGURE 2

Spreading resistance profiles after proton irradiations. Fluence: $1x10^{12}$ cm^{-2}.
2 MeV (#1). 3 MeV (#2), 5 MeV (#3)

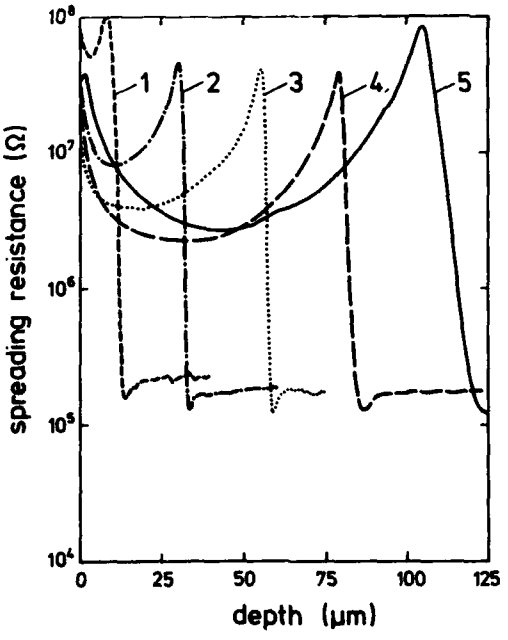


FIGURE 3

Spreading resistance profiles after α-particle irradiations. Fluence: $1x10^{12}cm^{-2}$.
3 MeV (#1), 6 MeV (#2), 8.8 MeV (#3), 11 MeV (#4), 13 MeV (#5)

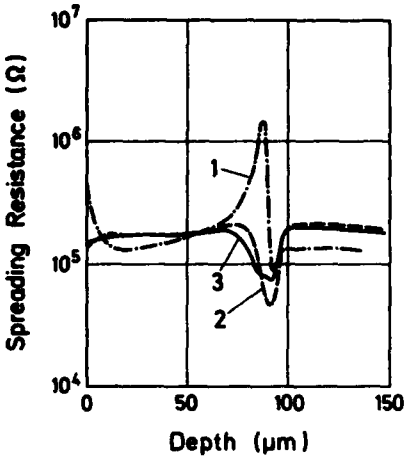


FIGURE 4

Spreading resistance profiles after thermal anneal. Irradiation: $1x10^{12}$ cm^{-2} H^{+}, 3 MeV.
Irradiated (#1), 350°C (#2), 400°C (#3)

The resistivity minimum corrosponds to the theoretical projected range, which depends on the particle energy.

After 350°C furnace annealing of proton-implanted silicon (fig. 4), a layer of increased conductivity evolves in the particle range due to the formation of hydrogen-donors [5]. Above 350°C, the doping concentration decreases.

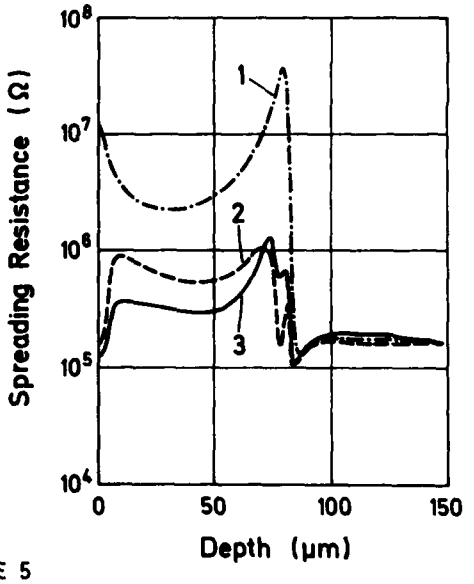


FIGURE 5

Spreading resistance profiles after thermal anneal. Irradiation: $1x10^{12}$ cm^{-2} He^{++}, 11 MeV
Irradiated (#1), 350°C (#2), 400°C (#3)

In α-particle irradiated silicon, no such donor formation was observed. After annealing at 350°C only a little "dip" in the resistivity profile is found (fig.5), which disappears at higher temperatures. This behaviour may be caused by energetically deep lying donors, reported by Yoshizawa et al [6].

3. RECOMBINATION CENTRE PROFILES

The recombination centres were investigated using DLTS. Proton irradiation induces a variety of defects (fig.6), most of them are not yet identified[7].

The defect spectrum after α-particle irradiation is very similar (fig.7), only the signal E(220K) in covered by an additional signal E(210K), and the minor peaks E(70K) and E(80K) are missing. The fraction of energetically deep lying levels, which are considered to be very effective recombination centres is higher than in the case of proton irradiations. The concentrations of all proton induced radiation defects show a marked peak at the particle range, Rp (fig.8). The defect concentration profiles after α-particle irradiation show the same behaviour (fig.9). The concentrations increase from the surface to the particle range and then drop rapidly. In this depth, the relative abandances of energetically deep lying levels are much higher, especially in the α-particle irradiated samples.

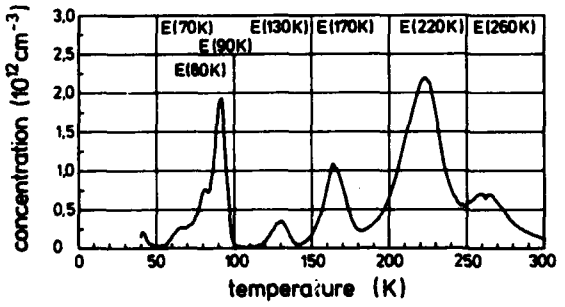


FIGURE 6

DLTS-Spectrum of irradiated n-type silicon.
Irradiation: 3×10^{10} cm^{-2} H^+, 3 MeV
Space charge layer just below the particle range (96-100 μm)

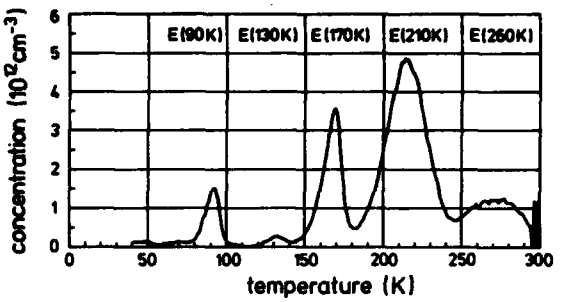


FIGURE 7

DLTS-Spectrum of irradiated n-type silicon.
Irradiation: 2×10^{10} cm^{-2} He^{++}, 8.8 MeV
Space charge layer just below the particle range (61-66 μm)

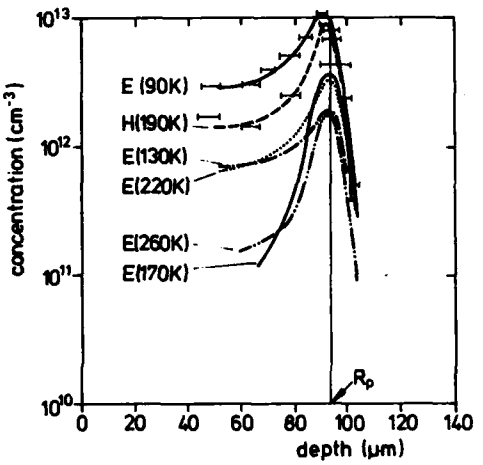


FIGURE 8

Defect profiles after 3 MeV H^+ irradiation (3×10^{10} cm^{-2}), not annealed

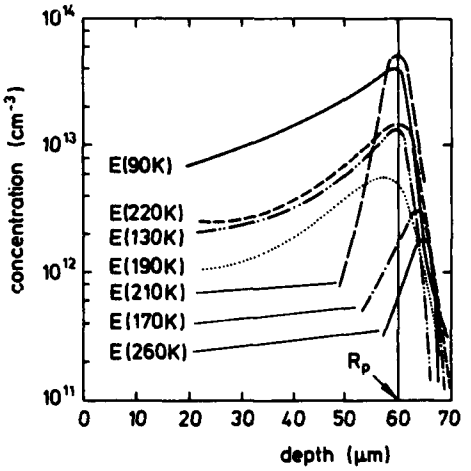


FIGURE 9

Defect profiles after 8.8 MeV He^{++} irradiation (2×10^{10} cm^{-2}), not annealed

4. CONSEQUENCES FOR POWER DEVICES

Proton irradiation is a versatile tool for localized lifetime control in semiconductor devices. The formation of hydrogen-donors in proton-irradiated devices (about $4x10^{14}$ cm^{-3} for $1x10^{12}cm^{-2}$ implant dose) inevitably leads to changes in the doping level and may thus affect the blocking capability of power devices in an undesired way.

Application of α-particle irradiation enables the creation of buried recombination layers without this doping effect.

The consequences are shown for power diodes:

Two groups of pnn^+-diodes with junction depth of 50 µm and 1360 V blocking voltage were irradiated with $3x10^{12}cm^{-2}$ protons at 2.25 MeV energy (corresponding to the example of fig.1) and with $3x10^{11}cm^{-2}$ α-particles at 8.8 MeV energy respectively, in order to create the maximum damage in about 10 µm below the junction in both cases. For thermal stabilisation the devices were annealed for 1/2h at 350°C. The forward voltage drop was 1.76 V (#1) and 1.63 V (#2), respectively at a current density of 110 A/cm^2. The very soft reverse recovery characteristics can be seen from fig.10. Whereas the blocking voltage of the α-particle irradiated diodes remained at 1360 V, the blocking voltage of the proton irradiated diodes degraded to 960 V.

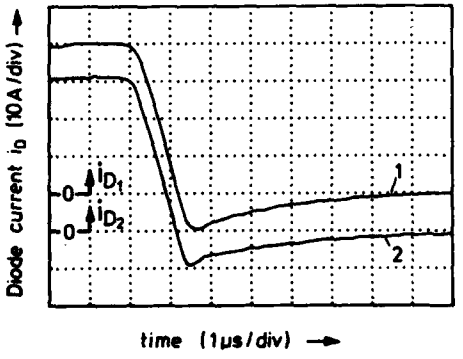


FIGURE 10

Reverse recovery of α-particle irradiated (1) and proton-irradiated (2) pnn^+-diodes with the same forward voltage drop. T_j=25°C

5. SUMMARY

Our results indicate, that α-particle irradiation is a promising technology for localized lifetime control in silicon devices. Because of the higher damage rate compared with proton irradiation, smaller particle fluences are necessary for a given switching behaviour. Furthermore, the fraction of deep recombination centres is remarkably larger in α-particle irradiated silicon.

This technology enables the creation of buried recombination layers without undesired donors. In this way, the dynamic properties of power diodes were optimized without sacrificing their blocking capabilities.

6. ACKNOWLEDGEMENTS

The irradiations were performed at the "Institut für Kernphysik", Frankfurt. The authors wish to thank W.-D.Nowak, D.Silber and B.Thomas for valuable discussions.

REFERENCES

1. D.Silber, W.-D.Nowak, W.Wondrak, B.Thomas and H.Berg
IEDM Technical Digest (1985), 162

2. A.Mogro-Campero, R.P. Love, M.F.Chang and R.F.Dyer
IEEE EDL-6(5), 224(1985)

3. T.Kushida, H.Tadano, S.Hashimoto, M.Takigawa, I.Igarashi and J.Nishizawa
Conf.Rec.IEEE Ind. Appl.Soc.1986, 372

4. W.Wondrak and D.Silber
Physica 129 B, 322 (1985)

5. Y.Ohmura, Y.Zohta and M.Kanazawa
Sol.St.Comm. 11,263 (1972)

6. M.Yoshizawa, M.Miyake and H.Harada
J.Electrochem.Soc.131(2), 453(1984)

7. K.Irmscher, H.Klose and K.Maass
J.Phys.C: Sol.St.Phys.17, 6317(1984)

Isolation Techniques in Power ICs with Vertical Current Flow

R. Zambrano, G. Ferla, S. Musumeci, M. Paparo

S G S Microelettronica s.p.a.
Stradale Primosole, 50 - 95121 Catania (Italy)

The increase in the range of applications of today's ICs is directly related to the fabrication of devices which can both withstand higher voltages and handle larger currents.

This can be accomplished either by improving the performances of the standard ICs, or by utilizing "intelligent" power transistors, e.g. monolithically integrating the power stage and a low voltage control circuitry.

To operate successfully in the high Volt*Ampere range, the power stage must have a vertical current flow and is therefore built on lightly doped n-type layers. It is possible to realize vertical power stages with either a single insulated collector (or drain) transistor, or with many transistors sharing a common collector (drain.)

The vertical current flow allows the best exploitation of the Si area, since it maximizes the current density, unlike the standard ICs, where the current flow makes a U-turn.

Figure 1 shows the difference between the two structures.

The technology described in this paper takes advantage of the better performances of the vertical structure transistors together with the well known low cost and design flexibility of standard bipolar ICs.

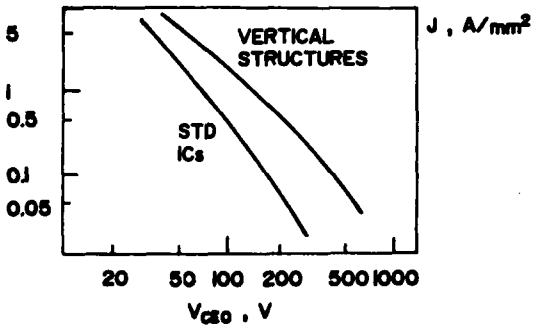


Figure 1

Current density vs. open base breakdown.

A p-type buried layer is realized to insulate the single components from each other and from the power part (fig. 2), this is not an easy task due to the fast "out-diffusion" of the boron towards the device surfaces.

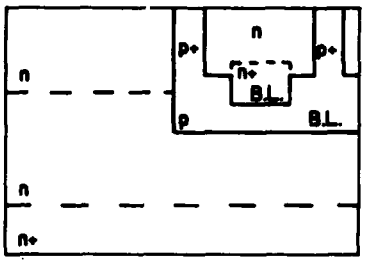


Figure 2

Vertical PIC isolation structure.

Figure 3 shows two dopant concentration profiles along a vertical section passing through the two buried layers.

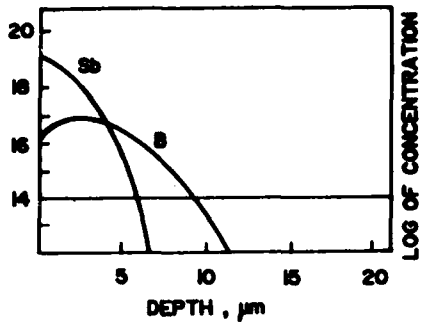


Figure 3a

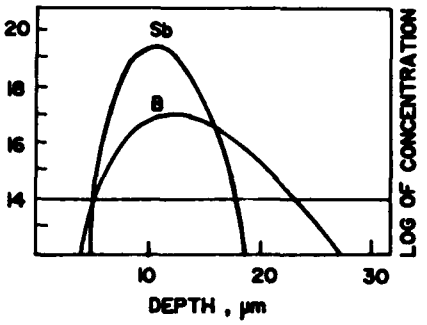


Figure 3b

Before the upper epitaxial growth the Sb (or As) and the B have already been diffused at high temperatures (3a), the ratio between the surface concentrations is large. After the epilayer growth and the isolation diffusion, however, the situation changes (3b.) The B diffuses out faster than Sb and at the interface between the n-type epilayer and the n-type buried layer their concentrations are similar.

This poses the problem that the boron concentration in the region between the surface and the n-type buried layer has to be smaller than the sum of the n-type impurities in order to avoid the appearance of phantom layers.

In addition the punch-through voltage of the p-type buried layer should be higher than the sum of the power and the low voltage transistor breakdowns and thus a bargain between these two requirements must be struck.

The solution cannot be found by increasing or decreasing the boron concentration; indeed, more sophisticate techniques have been developed.

Three solutions are possible, they can be utilized separately or in combination :

- increasing the concentration of the upper epitaxial layer,
- adding a lower dose of phosphorous in the n-type buried layer,
- placing an additional epilayer between the two buried layers.

Each of these techniques is aimed at a given voltage range : it is not possible to increase the concentration of the upper epilayer if a very high breakdown voltage is needed, nor is it possible to space apart the two buried layers if a low voltage device has to be made.

The evaluation of the junction isolation in standard ICs is done by considering three parameters, namely the current gain of the lateral npn parasitic transistor N1 (whose emitter, collector and base are two adjacent

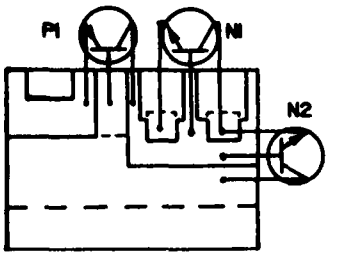


Figure 4
Parasitic transistors in PIC.

epitubs and the isolation region, respectively), its collector-base breakdown (shorted emitter), and its punch-through voltage.

In the PIC shown, two additional parasitic transistors (one npn and one pnp) must be taken into account. Their emitter, base and collector are, respectively :

1) the two buried layers and the n-type substrate (N2), and

2) the base of the power stage, the n-type substrate and the isolation region (P1.)

These three transistors are indicated in figure 4, figures 5 to 7 show their DC characteristics. Only the npn transistor whose emitter and base are

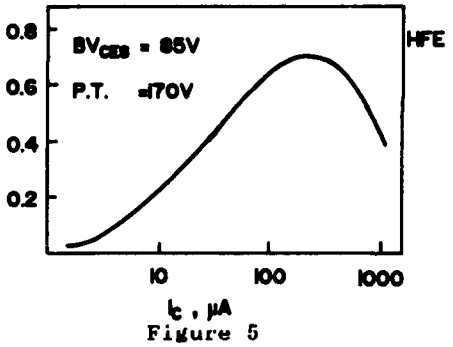


Figure 5

N1 current gain.

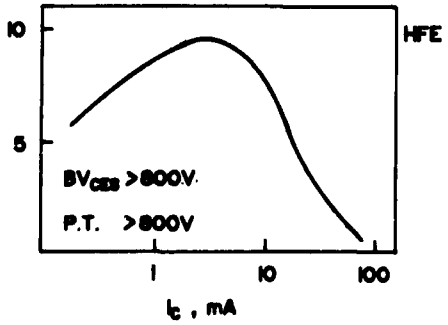


Figure 6

N2 current gain

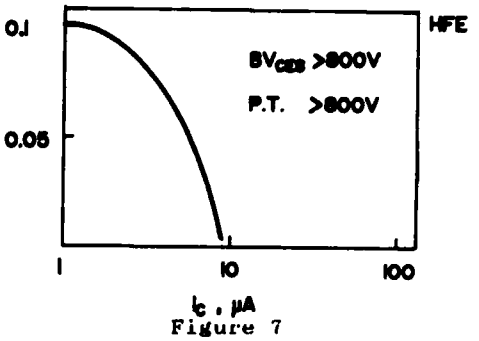


Figure 7

P1 current gain.

the two buried layers shows a current gain larger than the unity.

As previously explained the boron concentration cannot be increased, but it must be mentioned that the p-type buried layer is always reverse biased so that this npn transistor is normally off.



Figure 8

PIC die (26 sqmm.)

These isolation techniques have

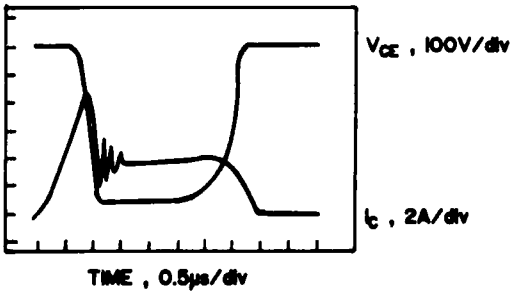


Figure 9

Collector voltage/current crossing.

already been implemented on pratical PIC devices having a darlington power stage and a control side driven by logic level inputs (fig. 8.)

A proper layout design of the device (optimized grounding network, power geometry) allows a 26 sqmm die to drive a 450 V / 5 A load (DC motor) with a switching speed in excess of 1000 V / microsec (fig. 9.)

Session P2.2

Posters

Tuesday, September 15, 1987

THE SMALL-SIGNAL BEHAVIOUR OF POLYCRYSTALLINE-SILICON MOSFET's

A. Gnudi, P. Ciampolini, R. Guerrieri, M. Rudan and G. Baccarani

Dipartimento di Elettronica, Università di Bologna
viale Risorgimento 2, 40136 Bologna, Italy

Abstract

In this paper we investigate the small-signal properties of polycrystalline-silicon MOSFET's by means of a newly-developed device-analysis program. The physical model of the grain-boundaries includes both donor and acceptor energy-distributed traps and accounts for time-dependent carrier-storage. The above phenomena are shown to be responsible for anomalously-large parasitic capacitances, which can seriously affect the transistor dynamic behaviour. Moreover, the admittance matrix turns out to be strongly non-reciprocal, and an equivalent circuit is identified characterized by a transadmittance with an imaginary part up to 3-4 orders of magnitude larger than the oxide admittance.

1. Introduction

Polycrystalline-silicon thin film transistors (TFT's) are currently investigated as active elements in liquid-crystal flat-panel displays and as a possible alternative to silicon on insulator (SOI) MOSFET's. In order to overcome the problems of TFT's, such as low carrier mobility, high leakage current and large threshold voltage, which are of major concern for dynamic applications, a large effort has been devoted to improve the quality of the film. Hydrogen passivation has resulted in reduced density of grain-boundary traps, while zone-melting recrystallization has been identified as a key process to increase the grain size. However, the overall electrical performances are not fully satisfactory, and are still essentially influenced by the large amount of traps at the grain boundaries.

The switching properties of TFT's play an important role when the drivers of the liquid-crystal transistor array are integrated onto the polycrystalline-silicon film. In this paper we investigate the dynamic behaviour of TFT's by using a two-dimensional device-analysis program, HFIELDS, developed at the University of Bologna. The program provides steady-state, small-signal and transient analysis, does not impose any restriction on device geometry, and can handle different semiconductor-semiconductor or semiconductor insulator interfaces, allowing for both donor and acceptor interface states. The latter, in turn, can be either monovalent or energy-distributed; thus the program can realistically simulate polycrystalline-silicon devices. In [1] it was shown that the simulation results by HFIELDS can qualitatively interpret experimentally observed phenomena such as the large threshold voltage and the transconductance degradation in subthreshold and strong inversion. The above result was achieved by idealizing the polycrystalline-silicon film by a number of equal-size grains, with strictly two-dimensional grain boundaries where a large amount of both donor and acceptor traps was accommodated. We believe that the same approach, even without trying to fit any experimental data, can help in understanding some remarkable capacitive effects which can negatively influence the dynamic behaviour of the device.

In the next section the small-signal model of the grain boundary is discussed, while section 3 is dedicated to the polycrystalline-silicon MOSFET simulation. Conclusions will follow in section 4.

2. The small-signal model of the grain boundary

The grain boundaries are treated as strictly two-dimensional interfaces, where both acceptor and donor traps with energy distributions $N_{it}^{(a)}(E_t)$, $N_{it}^{(d)}(E_t)$, respectively, are accommodated. The above states influence the electric potential distribution because of the trapped charge; on the other hand, according to Shockley-Read-Hall statistics, they enhance the generation-recombination rates, which, in turn, are related to the time dependent modulation of the trapped charge. By defining $n_{it} = n_{it}(E_t)$ the trapped-electron density at the level E_t, c_n and c_p the capture probabilities, e_n and e_p the emission probabilities of electrons and holes, respectively, the physical model is described by the following set of partial differential equations

$$\operatorname{div}\mathbf{D} = q(p - n + N) + \rho_{it}\,\delta(\mathbf{r} - \mathbf{r}_i) \qquad (1a)$$

$$\frac{\partial n}{\partial t} - \frac{1}{q}\operatorname{div}\mathbf{J}_n = (G - R)_{in}\,\delta(\mathbf{r} - \mathbf{r}_i) \qquad (1b)$$

$$\frac{\partial p}{\partial t} + \frac{1}{q}\operatorname{div}\mathbf{J}_p = (G - R)_{ip}\,\delta(\mathbf{r} - \mathbf{r}_i) \qquad (1c)$$

$$\frac{\partial n_{it}}{\partial t} = c_n n(N_{it} - n_{it}) - e_n n_{it} - c_p p n_{it} + e_p(N_{it} - n_{it}) \tag{1d}$$

comprising Poisson's equation (1a), carrier-continuity equations for electrons (1b) and holes (1c) and the detailed-balance equation for the trapped charge at the energy level E_t (1d). In (1d), $N_{it}(E_t) = N_{it}^{(a)}(E_t) + N_{it}^{(d)}(E_t)$, $e_n = c_n n_1$, $e_p = c_p p_1$, with $n_1 = n_i \exp[(E_t - E_i)/kT]$, $p_1 = n_i \exp[(E_i - E_t)/kT]$, E_i being the intrinsic Fermi level. Carrier transport is based on the drift diffusion mechanism, as in standard single-crystal devices. Among non-standard terms, ρ_{it} is the interface-trapped charge density per unit area, r_i denotes the i^{th} interface region and $(G-R)_{in}$, $(G-R)_{ip}$ are the net generation rates per unit area for electrons and holes, respectively. Remembering that the acceptor states are negatively charged when filled, and the donor states are positively charged when empty, we have

$$\rho_{it} = q \int_{E_v}^{E_c} [N_{it}^{(d)}(1 - f_{it}) - N_{it}^{(a)} f_{it}]\, dE_t \tag{2a}$$

$$(G-R)_{in} = \int_{E_v}^{E_c} c_n N_{it}[(n + n_1) f_{it} - n]\, dE_t \tag{2b}$$

$$(G-R)_{ip} = \int_{E_v}^{E_c} c_p N_{it}[p_1 - (p + p_1) f_{it}]\, dE_t, \tag{2c}$$

$f_{it}(E_t) = n_{it}(E_t)/N_{it}(E_t)$ being the occupation probability.

Notice that the above equations fully account for the time-dependent carrier storage effect. Integrating (1d) over the energy gap and using (2) we can write

$$\frac{1}{q}\frac{\partial \rho_{it}}{\partial t} = (G-R)_{in} - (G-R)_{ip}, \tag{3}$$

which clearly shows that only in steady-state the generation-recombination rates for electrons and holes become equal, and system (1) reduces to the standard form usually adopted by device-analysis programs.

The general solution of the above equations provides φ, n, p, f_{it} as a function of time over the device. The details of the numerical solution in the case of small-signal analysis are given in [2]. In the same work it was shown that, letting $\delta f_{it} = \mathrm{Re}[\tilde{f}_{it} \exp(j\omega t)]$ be the variation of f_{it} when a sinusoidal signal of angular frequency ω is applied to one electrode, the following expressions hold

$$\tilde{f}_{it} = [c_n(1 - f_{it}^{(o)})\tilde{n} - c_p f_{it}^{(o)} \tilde{p}]\, \frac{\tau(1 - j\omega\tau)}{1 + (\omega\tau)^2} \tag{4}$$

$$\tau = [c_n(n + n_1) + c_p(p + p_1)]^{-1}, \tag{5}$$

$f_{it}^{(o)}$ being the steady-state value given by

$$f_{it}^{(o)} = \frac{c_n n + c_p p_1}{c_n(n + n_1) + c_p(p + p_1)} \tag{6}$$

and $\tilde{n}$, $\tilde{p}$ the complex numbers representing the AC variations of the carrier concentrations; τ is the characteristic time-constant of the trapping-detrapping mechanism.

The dependence of $\tilde{f}_{it}$ on $\omega\tau$ is responsible for the anomalous capacitive behaviour of the grain boundary. In a former analysis of a silicon bicrystal, Seager and Pike identified an anomalous low-frequency capacitance effect due to the trapped-charge induced modulation of the barrier heigth which develops at the grain boundary under non-equilibrium conditions [3]. Similar results were obtained in [2] by using the above model of the grain boundary. Such an effect can be explained as follows. When a sinusoidal signal is applied to the bicrystal, an out-of-phase modulation of the trapped charge $\tilde{\rho}_{it} = -qN_{it}\tilde{f}_{it}$, with $\tilde{f}_{it}$ given by (4), follows. This induces an out-of-phase modulation of the potential barrier via Poisson's equation and since the current flowing across the barrier is an exponential function of the barrier height, its out-of-phase modulation gives rise to a large capacitive current proportional to the d.c. current. Under equilibrium conditions, due to symmetry reasons, no change of the trapped charge occurs, and the above effect vanishes; therefore only the capacitance term related to the depletion region around the interface is measurable. When the operating frequency is well above $1/\tau$, the trapped charge cannot follow any more the input signal, thus causing a dispersive effect of the capacitance curves.

The analytical model by Seager and Pike is based on the thermionic theory. However it has been shown that also the diffusion approach leads to the same exponential dependence of the d.c. current density on the barrier potential, only the pre-exponential factor being affected. Moreover, the diffusion approach is most appropriate for small barriers at low doping densities, when $kT/qE_{max} > \lambda$, E_{max} being the electric field at the interface and λ the mean free path of the carriers. Both models rely on some simplifying assumptions: uniformly doped crystals, abrupt depletion approximation, negligible minority carrier contribution to the current density. In TFT's the situation is even more complicated by the effect of the transverse electric field induced by the gate, and by the rapid variation of the electron concentration in the channel region normal to the semiconductor-oxide interface. The numerical approach followed here removes all these limitations, hence it is believed to provide reliable informations on the dynamic behaviour of the currently investigated devices.

3. The polycrystalline-silicon MOSFET

The mesh of the simulated polycrystalline-silicon MOSFET having a channel length $L = 6\ \mu m$ and a film thickness $t_s = 0.2\ \mu m$ is shown in figure 1; all the grains are $1\ \mu m$ long. The density of the interface states is $10^{12} cm^{-2}\, eV^{-1}$ at midgap and increases exponentially toward the band edges. This U-shaped energy distribution is consistent with several experimental results based on different techniques [4]. The distributed-line AC equivalent circuit is sketched in figure 2. Here, $G(V_G)$ and $C(V_G)$ represent the equivalent parallel conductance and capacitance associated with each grain boundary along the channel.

A small change of the gate voltage modulates the parallel capacitances $C(V_G)$, thereby generating an out of phase current flowing along the channel. The resulting capacitance C_{SG}, i.e. the current flowing out of the source contact following a change in gate voltage, turns out to be extremely large, and not related to the oxide capacitance. Thus, $C_{SG} >> C_{GS}$ (as the gate current must always flow across the distributed oxide capacitance) and a strongly non-reciprocal admittance matrix results. This conclusion is demonstrated in figures 3 and 4, where C_{GS}, C_{GD}, C_{SG} and C_{DG}, normalized to the oxide capacitance, are represented against the drain voltage for a given V_G. It should be noticed that C_{DG} is negative and nearly equal to $-C_{SG}$. The same figure also shows C_{SD} and C_{DS} which, as opposed to the standard MOSFET source-drain and drain-source capacitances, are both positive and large (up to some 1,000 times the oxide capacitance) as soon as the drain voltage exceeds approximately $1V$. As the capacitances associated with the boundaries are in series, the total capacitance is dominated by the smallest one. Finally, the oscillations in the curves are likely to be due to the small number of grains constituting the silicon thin film.

The lumped-element equivalent circuit is shown in figure 5, where C_{GS} , C_{GD} and C_{SD} are the same as before, G_0 is the equivalent conductance between drain and source, and the current generator $Y_m V_{GS}$ is the result of the non-reciprocity of the admittance matrix. Notice that $Y_m = G_m + j\omega C_m$, ω being the operating angular frequency, is a complex number representing the transadmittance of the MOSFET. Since we have only four linearly independent terms in the admittance matrix, given C_{SD}, C_{GD} and C_{GS}, the capacitance C_m is uniquely determined by one of the following relations

$$C_{DS} = C_m + C_{SD}$$

$$C_{SG} = C_m + C_{GS}$$

$$C_{GD} = C_m + C_{DG}$$

as can be inferred from the figure.

While C_{GS} and C_{GD} are strictly related to the oxide capacitance, all other components of the equivalent circuit essentially depend on the trapping-detrapping processes at the grain boundaries; therefore they display the same dispersive properties as the bicrystal when the operating frequency varies in the range of the audio frequencies. This is shown in figure 6, where it is evident that C_m and C_{SD} remain constant up to $1KHz$, for given values of the drain and gate voltages. For higher frequencies, both capacitances drop and the effect of the traps is no longer evident.

4. Conclusions

We have shown in this paper that the modulation of the barrier height which occurs at the grain boundaries of polycrystalline silicon MOSFET's as a consequence of the AC charge storage, is responsible for extremely large C_{DG} , C_{SG} and C_{DS} values, which negatively affect the dynamic behaviour of the above devices. The admittance matrix turns out to be strongly non-reciprocal. Such a behaviour can be fully described by an equivalent circuit which is characterized by the presence of the transadmittance $Y_m = G_m + j\omega C_m$, where C_m can be extremely high compared with the oxide capacitance. Such an effect, however, disappears at sufficiently-large operating frequencies, at which the trapping centers do not respond to the small-signal variations of the electric potential and carrier concentrations.

REFERENCES

[1] R. Guerrieri, P. Ciampolini, A. Gnudi, M. Gibertoni, M. Rudan and G. Baccarani: "An investigation on polycrystalline-silicon MOSFET operation", *Proceedings of the Second International Conference on Simulation of Semiconductor Devices and Processes*, Swansea, U.K, July 1986.

[2] A. Gnudi, P. Ciampolini, R. Guerrieri, M. Rudan, G. Baccarani: "Small-signal analysis of semiconductor devices containing generation-recombination centers", *NASECODE V Conf.*, June 1987, Dublin.

[3] C. H. Seager and G. E. Pike: "Anomalous low-frequency grain-boundary capacitance in silicon", *Appl. Phys. Lett.*, vol. 37, pp. 747-749, 1980.

[4] H. C. de Graaff, M. Huybers and J. G. de Groot: "Grain boundary states and the characteristics of lateral polysilicon diodes", *Solide-State Electronics*, vol. 25, no. 1, p. 67-71, 1982, and references therein.

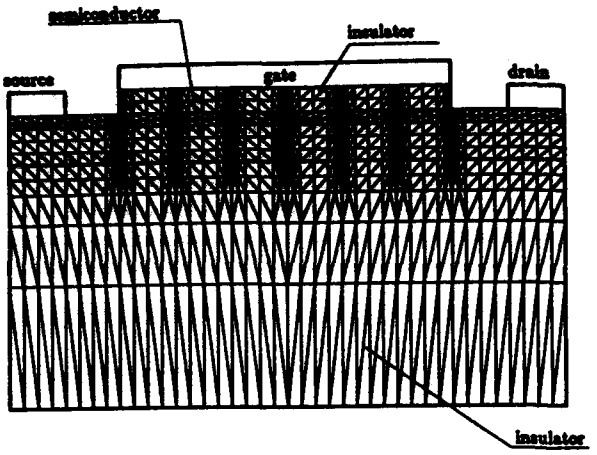


Fig. 1: Mesh of the simulated polycrystalline-silicon MOSFET.

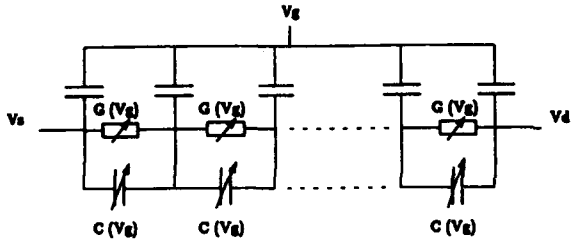


Fig. 2: Equivalent circuit of the polycrystalline-silicon MOSFET under non-equilibrium conditions.

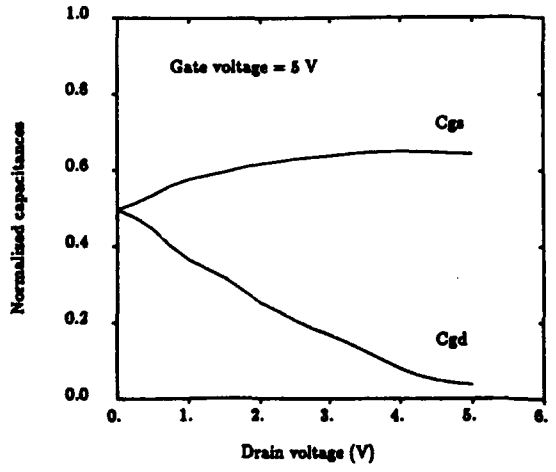


Fig. 3: Normalized capacitances $C_{gd} = C_{GD}/C_{ox}$ and $C_{gs} = C_{GS}/C_{ox}$ vs the drain voltage.

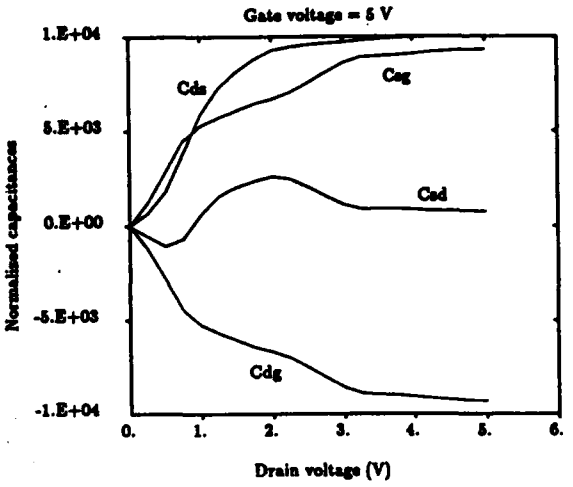


Fig. 4: Normalized capacitances C_{dg}, C_{sg}, C_{ds}, C_{sd} vs the drain voltage.

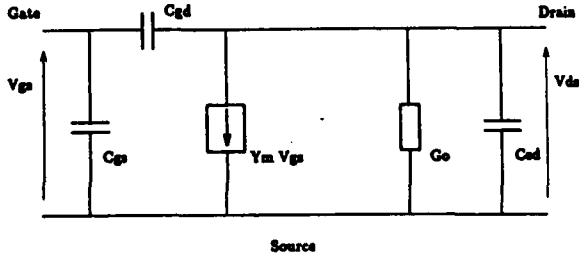


Fig. 5: The lumped-element equivalent circuit of the polycrystalline-silicon MOSFET. $Y_m = g_m + j\omega C_m$ is the transistor transadmittance.

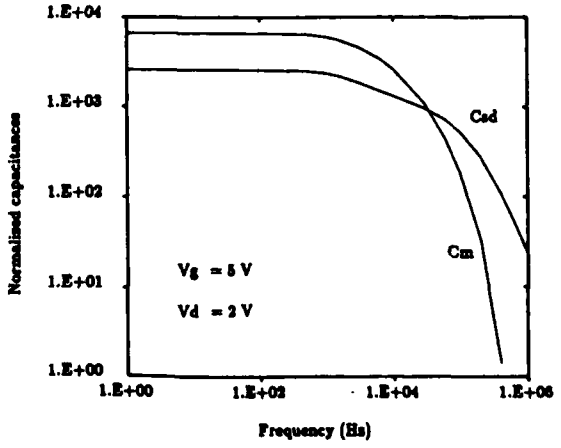


Fig. 6: Normalized capacitances C_{sd} and C_m vs frequency.

P2.2.2

IMPROVED DETERMINATION OF SURFACE MOBILITY AT MOSFETS WITH THIN GATE OXIDE

W. Soppa, H.-G. Wagemann

Technische Universität Berlin
Institut für Werkstoffe der Elektrotechnik
Jebensstr. 1, D-1000 Berlin 12, West-Germany

We propose an extraction procedure to determine the surface mobility of MOSFETs based on the drain current equation after Pao and Sah. This model is extended to consider short channel effects and the charge of interface states in nonequilibrium. At low drain voltage the calculated drain current is compared to two-dimensional simulation results. Excellent agreement is found for samples with oxide thickness down to 20 nm even for MOSFETs with channel length of submicron dimensions.

Extracted mobility values are in good agreement with experimental results of time-of-flight measurements and are found to be independent from oxide thickness. At low oxide thickness the surface mobility extraction using the charge sheet model leads to remarkable different values.

1. INTRODUCTION

Electrical characterization of scaled semiconductor devices is indispensible for the development of VLSI circuits and technology. One of the most interesting quantities for MOSFETs is the drain current as function of gate and drain bias. Its calculation requires a reliable mobility model which should be able to consider all important influences on surface charge transport: doping as well as surface roughness, parasitic oxide and interface charges. For the evaluation of such a model it is essential to get meaningful experimental results for different processing conditions.

A widely used extraction procedure is calculating the surface mobility from measured MOSFET dc transfer characteristics by using the charge sheet drain current model (see e.g.[1]). It has been shown [2] that this model fails at low gate fields near threshold especially when analyzing devices with thin gate oxide. This is due to assuming the gate capacitance equaling the constant oxide capacitance C_{ox} which is too strong a simplification.

In this paper we present an improved extraction procedure for evaluating the surface mobility at MOSFETs especially with thin gate oxide. It is based on the extended drain current model of Pao and Sah [3] which is compared to two-dimensional simulation.

Results from n-channel MOSFETs fabricated under different processing conditions are presented and compared to experimental data of time-of-flight measurements. The difference in extracted values of surface mobility compared to those from the standard extraction procedure are discussed.

2. EXTENDED DRAIN CURRENT MODEL

Pao and Sah [3] have formulated the drain current of a MOSFET in the gradual channel approximation. By defining the effective surface mobility averaged over all channel carriers n

$$\mu_{eff} = \frac{\int_0^{x_c} \mu(x)\cdot n(x)\,dx}{\int_0^{x_c} n(x)\,dx} = \frac{\int_{\varphi_S}^{\varphi_B} \frac{\mu(\varphi)\cdot n(\varphi)}{\frac{d\varphi}{dx}}\,d\varphi}{\int_{\varphi_S}^{\varphi_B} \frac{n(\varphi)}{\frac{d\varphi}{dx}}\,d\varphi} \qquad (1)$$

with φ being the semiconductor potential, and restricting to the case of low fields parallel to the silicon-silicondioxide interface (that means low drain voltage), the drain current may be written as

$$I_D = \frac{Z}{L}\, q\, n_i \beta L_{Di} \mu_{eff} \int_0^{U_{DS}} \int_{\varphi_S}^{\varphi_B} \frac{e^{\beta(\varphi-\xi)}}{F(\varphi,\xi)}\, d\varphi d\xi \quad . \quad (2)$$

L_{Di} is the Debye length of the intrinsic semiconductor, $\beta = q/kT$, φ_B is the bulk potential and Z, L, q and n_i have their common meanings. The nonequilibrium F-function is given by

$$F(\varphi,\xi) = -\,\mathrm{sign}(\varphi-\varphi_B)\cdot\{e^{\beta(\varphi-\xi)} - e^{\beta(\varphi_B-\xi)} + e^{-\beta\varphi} - e^{-\beta\varphi_B} + F_{KK}^2 \cdot \beta(\varphi-\varphi_B)[e^{-\beta\varphi_B} - e^{\beta\varphi_B}]\}^{1/2} \quad . \quad (3)$$

F_{KK} represents the short channel factor and will be discussed later, for long channel devices $F_{KK} = 1$. The nonequilibrium potential ξ is determined by the difference between hole and electron quasi fermi potential. The surface potential φ_S is determined by the voltage drop across the MOS system

$$U_{GS} = \varphi_S - \varphi_B - \frac{1}{C_{ox}}\,[Q_{sc}(\varphi_S,\xi) + Q_f + Q_{it}(\varphi_S,\xi)] + \frac{\Phi_{MS}}{q} \quad (4)$$

Q_f is the fixed charge, Q_{it} the charge of interface states and Φ_{MS} the work function difference. These quantities are frequently combined to the flat band voltage.

For the purpose of modeling real MOSFETs properly, two important extensions are required. The first one is a model for the charge Q_{it} of interface states in stationary nonequilibrium. As has been found [4], [5], donor type interface states D_{it_D} are mainly located in the lower half and acceptor type interface states D_{it_A} mainly in the upper half of the silicon bandgap. Therefore, the charge of interface states is given by

$$Q_{it}(\varphi) = q^2 \int_{-\infty}^{\infty} D_{it_D}(\varphi)\, f_p(\varphi)\, d\varphi - q^2 \int_{-\infty}^{\infty} D_{it_A}(\varphi)\cdot f_n(\varphi)\, d\varphi \quad . \quad (5)$$

$f_{n,p}$ is the occupation probability of traps with electrons or holes, respectively [6]. The density $D_{it}(\varphi)$ of interface states in the silicon bandgap is determined experimentally by analyzing quasistatic capacitance of MOS-varactors and is empirically approximated [5], [7] by

$$D_{it}(\varphi_S) = D_{it_o} + D_{it_1} \cosh\left(\frac{\varphi_S}{\varphi_o}\right) \quad . \quad (6)$$

D_{it_o}, D_{it_1} and φ_o are fitted to experimental data. As long as $D_{it}(\varphi)$ is a slowly varying function, the following approximation of eq.(5) holds

$$Q_{it}(\varphi,\xi) = q^2 \int_{\varphi_S}^{o} D_{it_D}(\varphi)\, d\varphi - q^2 \int_{o}^{\varphi_S-\xi} D_{it_A}(\varphi)\, d\varphi \quad (7)$$

This is shown in fig. 1.

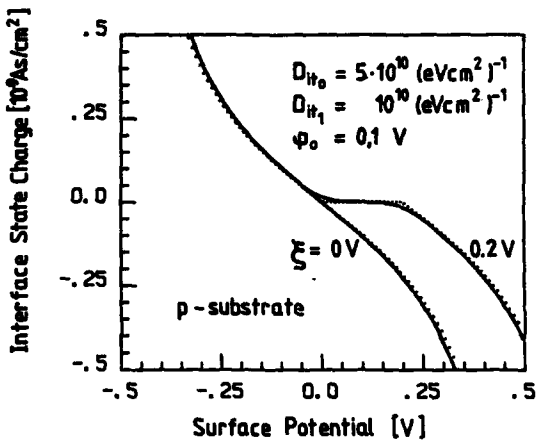


FIGURE 1

Charge of interface states versus surface potential calculated after eq. (5) (solid line and eq. (7) (dotted line) for $D_{it}(\varphi)$ after eq. (6).

The second extension of the model deals with short channel effects. The charge sharing effect [8] is considered in the calculation of the F-function (eq.(3)) by introducing the short channel factor F_{KK} which is calculated after Fichtner and Pötzl [9]

$$F_{KK} = 1 - \frac{x_j}{2L}\left[\sqrt{1+\frac{2x_S}{x_j}} - 1 + \sqrt{1+\frac{2x_S}{x_j}} - 1\right],$$

$$x_S = \sqrt{\frac{2\varepsilon_o\varepsilon_{Si}}{qN_A}\,(U_D - \varphi_S + \varphi_B)}\quad , \quad (8)$$

$$x_d = \sqrt{\frac{2\varepsilon_o\varepsilon_{Si}}{qN_A}\,(U_D - \varphi_S + \varphi_B + U_{DS})}\quad .$$

U_D is the built-in potential and x_j the source and drain junction depth. As shown in fig. 2, the charge sharing effect influences the transfer characteristics of a MOSFET mainly in weak inversion, because in strong inversion the ionized impurities play a minor role in determining the charge balance and field distribution in the device.

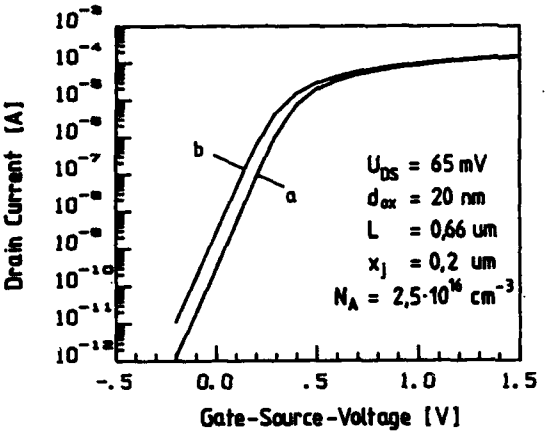


FIGURE 2

MOSFET transfer characteristics calculated after the Pao-Sah model (line a) and by considering the charge sharing effect eq. (8) (line b).

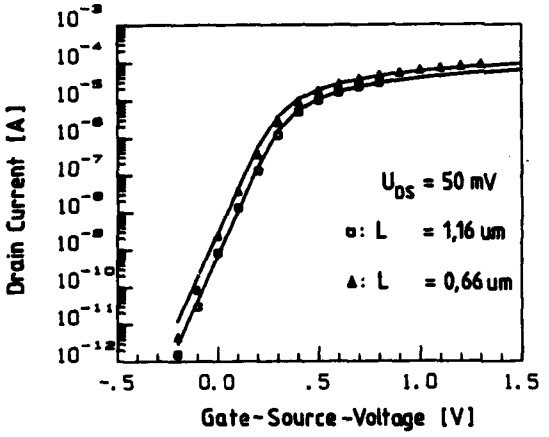


FIGURE 3

Extended Pao-Sah model (solid line) compared to results of MINIMOS (symbols) for devices with geometries of fig. 2.

In fig. 3 the extended Pao-Sah model is compared to two-dimensional calculations done by the device simulator MINIMOS for different channel lengths. In both models the surface mobility is set to the bulk value in order to suppress differences in drain current caused by different mobility models. Excellent agreement is found for devices down to submicron dimensions of channel lengths.

3. EXPERIMENTAL RESULTS OF EXTRACTED SURFACE MOBILITY

NMOS and PMOS samples have been fabricated in standard Al-gate and poly-silicon gate technology with typical device dimensions of 100 nm oxide thickness and 2 µm junction depth as well as in scaled poly-silicon gate CMOS technology with 0.3 µm junction depth and 36 nm or 20 nm oxide thickness, respectively. MOSFETs with varying channel lengths of 1 to 20 µm were used to determine the effective channel length and series resistance [10] and a large geometry MOSFET and MOS varactor with Z x L =(600 µm)² were used for capacitance measurements and determination of oxide charge Q_f and density D_{it} of interface states. The doping profile was determined from pulsed capacitor measurements and backgate bias variation.

The surface mobility was determined from measured transfer characteristics of MOSFETs with low Q_f and D_{it} at low drain voltage using the extended Pao-Sah model. Fig. 4 shows results plotted versus the effective gate field as introduced by Sabnis and Clemens [11] for devices with different values of oxide thickness. The experimental data are very similar to those determined from time-of-flight measurements [12] even for devices with 20 nm oxide thickness.

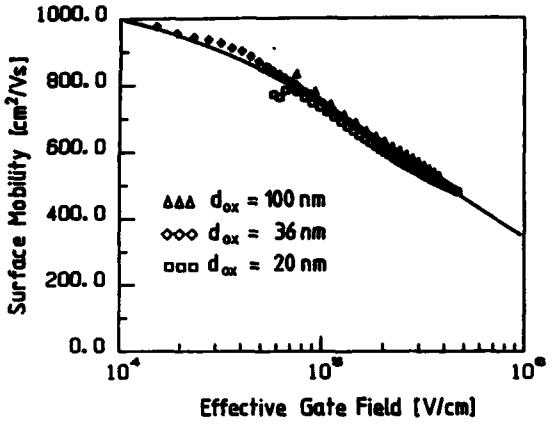


FIGURE 4

Extracted surface mobility for MOSFETs with different values of oxide thickness (symbols) compared to time-of-flight data (solid line)

A comparison of the presented extraction procedure to the standard method using the charge sheet model is shown in fig. 5. While at sufficient high gate fields the evaluated mobility values are the same, at lower fields near threshold a remarkable discrepancy can be realized. This difference becomes evident at low oxide thicknesses. In this case, the standard extraction procedure may lead to erroneous experimental data for the surface mobility.

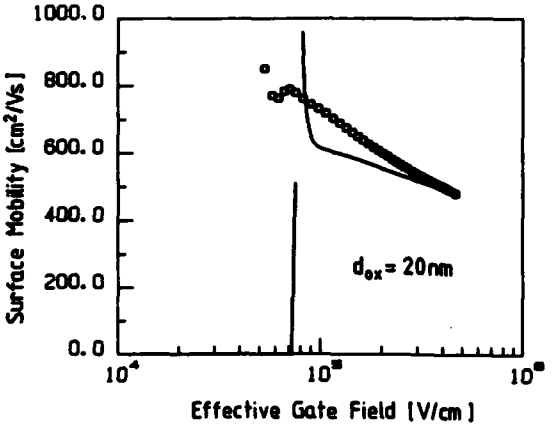


FIGURE 5

Extracted surface mobility data determined by the extended Pao-Sah model (symbols) or standard extraction procedure (solid line).

4. CONCLUSIONS

The Pao-Sah model extended to consider charge of interface states and short channel effects is useful for extracting the surface mobility from measured MOSFET transfer characteristics. The results are comparable to time-of-flight measurements. Even for short channel devices meaningful surface mobility values can be determined, as shown by comparison to two-dimensional simulation.

The extraction procedure leads to results which differ from those determined by use of the charge-sheet model especially at low oxide thickness. As a result of applying this improved procedure it is found that mobility data evaluated from MOSFETs with varying oxide thickness do not depend on this parameter.

REFERENCES

[1] Sze, S.M., Physics of semiconductor devices, Wiley & Sons, (1981) 440

[2] Sodini, C.G., Eckstedt, T.W., Moll, J.L., Solid-State Electr. 25, (1982) 833-841

[3] Pao, H.C., Sah, C.T., Solid-State Electr. 9 (1966) 927-937

[4] Knoll, M., Bräunig, D., Fahrner, W.R., IEEE Trans. NS-29, (1982) 1471-1482

[5] Kämpf, U., Wagemann, H.-G., IEEE Trans. ED-23, (1976) 5-10

[6] Shockley, W., Read, Jr. W.T., Phys. Rev. 87 (1952), 835-842

[7] White, M.H., Cricchi, J.R., IEEE Trans. ED-19, (1972), 1280-1288

[8] Yau, L.D., Solid-State Electr. 17 (1974), 1059-1063

[9] Fichtner, W., Pötzl, H.W., Int. Journ. Electron. 46 (1979), 33-55

[10] Terada, K., Muta, H., Jap. Journ. Appl. Phys. (1979), 953-959

[11] Sabnis, A.G., Clemens, J.T., IEEE Tech. Dig. IEDM (1979), 18-21

[12] Cooper, Jr. J.A., Nelson, D.F., Journ. Appl. Phys. 54 (1983), 1445-1456

Mobility Model for Silicon Inversion Layers

A.J. Walker and P.H. Woerlee

Philips Research Laboratories
5600 JA Eindhoven The Netherlands

The room temperature low field mobility of electrons and holes in silicon inversion layers has been studied to improve the mobility modeling. Samples with gate oxide thicknesses between 10nm and 50nm and surface doping levels of up to $4.5 \times 10^{17} cm^{-3}$ have been used. Three scattering mechanisms were considered and an accurate mobility model, derived from these, will be presented.

INTRODUCTION

The fabrication of sub-micron metal-oxide-silicon (MOS) transistors has become of major importance in recent years. Consequent problems have arisen in attempting to model the mobility of the charge carriers in the channels of such devices. In many instances, empirical models have been used and have been effective in describing many devices. However, the advent of sub-micron transistors has shown the presence of certain weaknesses in these models.

This paper presents a physically-based mobility model for the carriers in the inversion layer of MOS transistors derived from considering three mechanisms involved in the scattering of charge carriers in a surface inversion layer. As has been done elsewhere [3], both the resultant derived and measured mobilities are presented as functions of the effective normal field experienced by the carriers, a method which has led elsewhere [1] to an empirical relation known as the "universal curve" which is independent of oxide thickness and surface doping density. This relation is defined by:

$$\mu = \mu_{max}(\frac{E_{eff}}{E_c})^{C1} \qquad (1)$$

with C1 in the range between -0.20 and -0.33. E_c and μ_{max} are parameters, values of which can be found in [1].

Plots of this curve have been made in Figs 1 and 2 along with the measured mobility from certain devices to be described later. As can be seen, the mobility fall-off at high fields is not followed by equation 1. Unfortunately, this is precisely the working region for MOS transistors with gate oxide thicknesses in the neighbourhood of 10nm and with high gate drive.

MOBILITY

The clear aim of this study was to obtain a model for the low source-drain field mobility which could describe the mobility over a larger range of the effective field than could equation 1. As has been done elsewhere [5], the mobility was derived by considering three mechanisms which are believed to dominate the scattering of charge carriers in inversion layers. These were (i) Coulomb scattering, (ii) Electron - phonon scattering and (iii) Surface roughness scattering.

In this analysis, the following notation is used: T is the absolute temperature, Q_N is the mobile channel charge per unit surface area, Q_B is the bulk charge per unit surface area, a and b are constants, and ϵ_s is the permittivity of silicon.

For simplicity, a constant value, μ_0, is used for the Coulomb scattering. For electron - phonon scattering, the prediction from theory (see [2] and references therein) is :

$$\mu_{ep} \approx aT^{-1}(Q_N/3 + Q_B)^{-\frac{1}{3}} \qquad (2)$$

Surface roughness scattering, according to theory (see references in [2]), obeys the following relation :

$$\mu_{sr} = b(Q_N/2 + Q_B)^{-2} \qquad (3)$$

As has been previously stated, the mobility is presented as a function of the effective normal electric field E_{eff} experienced by the carriers in the inversion layer. The dependence of E_{eff} on surface charge can be estimated by the use

of Gauss's Law in one dimension, giving

$$E_{eff} = \frac{1}{\epsilon_s}(\frac{Q_N}{2} + Q_B) \quad (4)$$

With equation 4 in mind, the carrier-phonon and surface roughness mobilities can be approximately expressed respectively as:

$$\mu_{ep} = k_2 E_{eff}^{-1/3} \quad (5)$$

$$\mu_{sr} = k_1 E_{eff}^{-2} \quad (6)$$

where k_1 and k_2 are constants with respect to E_{eff}.

Matthiesen's rule was used to combine the three mobilities. i.e.

$$\frac{1}{\mu} = \frac{1}{\mu_0} + \frac{1}{\mu_{ep}} + \frac{1}{\mu_{sr}} \quad (7)$$

This gave the net mobility which can be expressed as follows:

$$\mu = \frac{\mu_0}{1 + \Theta_1 E_{eff}^2 + \Theta_2 E_{eff}^{0.33}} \quad (8)$$

where μ_0, Θ_1 and Θ_2 are three fitting parameters, since k_1 and k_2 cannot be derived theoretically. These can be determined from a fit of equation 8 to experimental data. Note that $\Theta_1 = \mu_0/k_1$ and $\Theta_2 = \mu_0/k_2$.

EXPERIMENTAL PROCEDURE

Measurements in this experiment were carried out on large (20 μm x 20 μm) MOS transistors with n-type polysilicon gates, and large (200 μm x 200 μm) MOS capacitors. The substrate material was 20 Ω-cm p-type 100 silicon for NMOS and 12 Ω-cm n-type 100 silicon for PMOS. For the NMOS case, the gate oxide thicknesses in the range between 20nm and 50nm (series A) were grown in dry oxygen at 950^oC while those with 10nm (series B) were grown at 900^oC in a diluted oxygen/nitrogen mixture (10% oxygen by volume). For the PMOS case, all the gate oxides were of 10nm thickness and were grown as for the series B NMOS samples.

All samples were given a post-metallisation anneal in forming gas at 450^oC for 30 minutes. A summary of the samples used can be seen in the table which includes data to be explained later. N_{surf} is the surface dope concentration in the channel region determined from using the simulation program SUPREM3.

The method used to experimentally extract the mobility can be found in the paper by Sodini et.al [4].

Type	N_{surf} (cm^{-3})	t_{ox} (nm)	μ_0 ($cm^2V^{-1}s^{-1}$)	k_1 Vs^{-1}
NMOS A	5×10^{16}	20	1659	5.99×10^{14}
NMOS A	5×10^{16}	30	1478	6.80×10^{14}
NMOS A	5×10^{16}	40	1492	7.15×10^{14}
NMOS A	5×10^{16}	50	1467	9.01×10^{14}
NMOS B	1×10^{17}	10	1227	5.51×10^{14}
NMOS B	2×10^{17}	10	998	7.40×10^{14}
NMOS B	3.5×10^{17}	10	1137	5.93×10^{14}
PMOS	1×10^{17}	10	451	2.01×10^{14}
PMOS	2×10^{17}	10	457	2.29×10^{14}
PMOS	1×10^{17}	10	296	2.12×10^{14}

Measurements were carried out using an HP 4140B pA meter/dc voltage source with the samples housed in a light-tight, electrically-shielded cabinet.

RESULTS

Fig.1 shows the results of measurements carried out on the series A NMOS samples. For comparison, the "universal" curve (equation 1) has also been drawn in. As can be seen, this empirical relation fits the data well for effective fields below $6 \times 10^5 V cm^{-1}$ but cannot follow the mobility degradation at higher fields. Fig.2 shows measurements made on the series B NMOS samples. As before, equation 1 has been drawn in and is seen to fit the data for effective fields below $6 \times 10^5 V cm^{-1}$ but again does not reflect the mobility degradation at higher fields.

Fig.3 shows fits of equation 8 to data from two NMOS samples with different oxide thicknesses and surface doping densities. In this case, it can be seen that the new physically-based model fits the measured data over a much wider range than does equation 1. That is, the mobility degradation at high fields is incorporated into the new relation. It was found that relative deviations between the fit and the data were smaller than 2 percent i.e. within the measurement accuracy. Note that the fits in Fig.3 differ slightly even for fields below $6 \times 10^5 V/cm$. This may be due to the second sample having a much larger surface doping density and thus a smaller mobility in this region due to Coulomb scattering.

The weak dependence of the mobility on E_{eff} observed for $E_{eff} < 6 \times 10^5 V cm^{-1}$ indicates that phonon scattering is

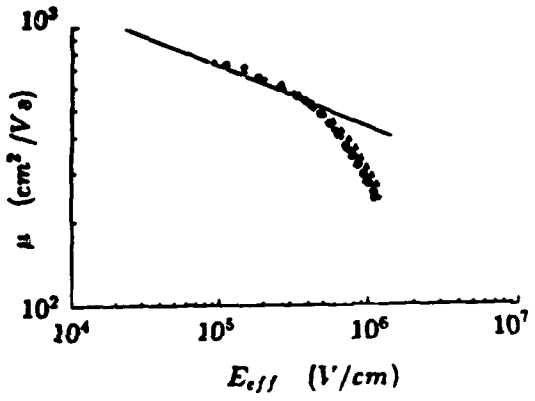


Fig.1 : The measured mobility with gate oxides in the range 20 nm - 50 nm (series NMOS A). The solid line is the universal curve. $E_c = 2.6 \times 10^4 V/cm$, $\mu_{max} = 927 cm^2/Vs$, $C_1 = 0.22$
($\diamond$: $20nm$, $\bigcirc$: $30nm$, $+$: $40nm$, $\triangle$: $50nm$)

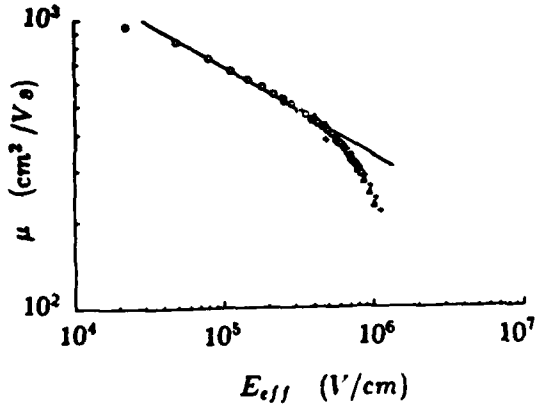


Fig.2 : The measured mobility with gate oxides of 10 nm (series NMOS B). The solid line is the universal curve with $E_c = 2.6 \times 10^4 V/cm$, $\mu_{max} = 1010 cm^2/Vs$, $C_1 = 0.31$
The surface dope concentrations are as follows:
$\bigcirc$: $10^{15} cm^{-3}$, $\triangle$: $2 \times 10^{17} cm^{-3}$, $+$: $4.5 \times 10^{17} cm^{-3}$

dominant in that region. It can be said that the magnitude of the scattering in this phonon-dominated region depends only on the charge distribution in the inversion layer and not on the details of the oxide growth.

The stronger decrease in the mobility for $E_{eff} > 6 \times 10^5 V/cm$ indicates that surface roughness scattering becomes more influential in this region. In this region of E_{eff}, the details of the oxidation process used to grow the gate may

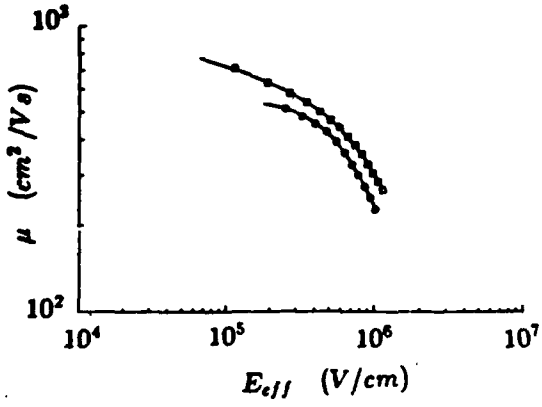


Fig.3 : Fits of new mobility model to measured NMOS data. The solid lines are the fits and the symbols, the data.
($\square$: $t_{ox} = 50nm, n_{surf} = 5 \times 10^{16} cm^{-3}$; $\bigcirc$: $t_{ox} = 10nm, n_{surf} = 2 \times 10^{17} cm^{-3}$)

influence the mobility. In the measurements carried out for this paper, there was a systematic decrease in mobility with decreasing oxide thickness. This could be tentatively explained by assuming that the thinner gate oxides have more surface roughness.

An interesting point to note is that equation 8 could be fitted to the results from different samples by taking a constant value of Θ_2 i.e. a constant value for k_2 for all the samples. The value for k_2 for the NMOS samples was different to that for the PMOS samples. If k_2 was fixed at $6.5 \times 10^4 cm^{5/3} V^{-2/3} s^{-1}$ for the NMOS samples and $1.3 \times 10^4 cm^{5/3} V^{-2/3} s^{-1}$ for the PMOS samples, then the values for μ_0 and k_1 needed for reasonable fits over the measured E_{eff} range can be found in the table. Fig.4 shows a fit of equation 8 to PMOS data from the last sample in the table with the values for the parameters given. This suggests that carrier - phonon scattering gives rise to the "universal curve" type of expression as in equation 1 and is independent of the device except for it being N or P type. Fig.5 shows plots of equation 8 for a typical NMOS and a typical PMOS with values for the parameters given in the caption.

CONCLUSIONS

In summary, it can be seen that a fairly simple physically-

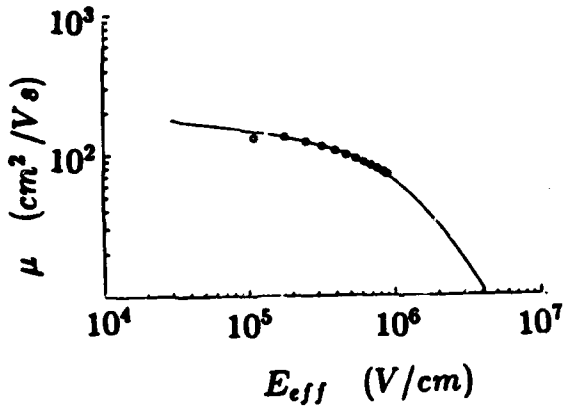


Fig.4 : Fit of new mobility model to measured PMOS data from last sample in table.

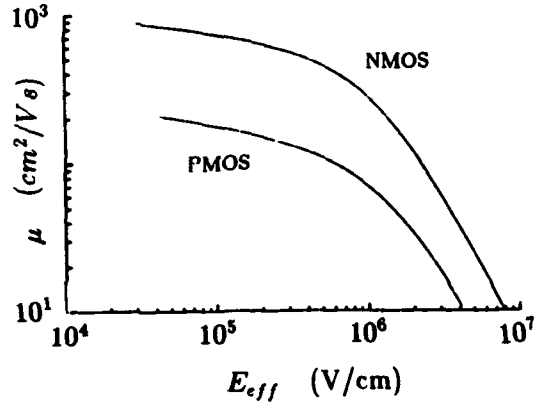


Fig.5 : Plots of equation 8 for typical NMOS and PMOS.
NMOS : μ_0=1480, k_1=6.5 x 10^{14}, k_2=6.5 x 10^4.
PMOS : μ_0=450, k_1=2.1 x 10^{14}, k_2=1.3 x 10^4.
The following units are used: μ_0 in cm^2/Vs, k_1 in V/s and k_2 in $cm^{5/3}/V^{2/3}s$

based mobility model, comprising three scattering mechanisms, can be used to describe electron and hole mobility data over a wide range of oxide thicknesses and surface doping densities. Furthermore, at effective fields below $6 \times 10^5 V/cm$, Coulomb and electron-phonon scattering are the dominant mechanisms while, at higher fields, electron-phonon and surface roughness scattering prevail. A simple model obtained from theoretical expressions for these mechanisms has been shown to be very accurate for a wide variety of samples.

REFERENCES

1. S.C. Sun and J.D. Plummer, IEEE Trans. Electron Dev., ED-27 1497 (1980).

2. T. Ando, A.B Fowler and F. Stern, Rev. of Modern Phys., 54 437 (1982).

3. A.G. Sabnis and J.T. Clemens, IEDM Tech. Dig., 18 (1979).

4. C.G Sodini, T.W. Ekstedt and J.L. Moll, Solid State Electronics 25, 833 (1982).

5. A. Hartstein, A.B. Fowler and M. Albert, Surface Science 98 181 (1980).

6. Y.C.Cheng and E.A. Sullivan, J. Appl. Phys., Vol.44, No.2, 923 (1973)

LEAKAGE CURRENTS IN LOW TEMPERATURE PROCESSED POLYCRYSTALLINE-SI TFTs

S.D. BROTHERTON, N.D. YOUNG, A. GILL

Philips Research Laboratories, Redhill, Surrey, U.K.

The influence of plasma hydrogenation on the leakage current of n-channel poly-Si TFTs has been examined. With low temperature processed devices, the major effect has been the suppression of gate controlled hole currents as a result of the increase in carrier generation lifetime. This has also resulted in a complementary increase in optical sensitivity.

1. INTRODUCTION

There is presently considerable interest in the performance of IGFETs fabricated from poly-crystalline silicon both for 3-D integrated circuit applications [1] and for addressing elements in liquid crystal displays [2]. The technologies investigated for these applications have differed principally in the technique used for the gate dielectric. In the former case it is usually grown by conventional thermal oxidation whereas for the latter application both high temperature growth and low temperature deposition have been investigated. One common feature in most of this work has been the use of plasma hydrogenation to passivate grain boundary trapping states and thereby improve the transistors on-characteristics. In the work presented below, whilst illustrating the improvement in the on-characteristic of low temperature processed TFTs, we will be concentrating upon the effect of the hydrogenation on the off-characteristic or device leakage current.

2. SAMPLE FABRICATION

The devices used in this work were auto-registered n-channel poly-Si TFTs fabricated with a low temperature technology (<620°C) and thus employing a deposited gate dielectric. The thickness of the poly-Si layer used in the transistor body was 0.2μm and the source and drain regions were doped by phosphorus ion implantation. The gate dielectric thickness was 0.15μm and the TFT channel width was 50μm with 10μm channel length.

3. RESULTS

The poly-Si grain diameter was ~700Å and given that the poly-Si was undoped it is fully depleted in the absence of any externally applied potential, with an intergrain potential barrier of <<kT. The conventional model of grain boundary traps assumes a neutral level near mid-gap which would make the material intrinsic. In confirmation of this, the activation energy of the leakage current minimum has been measured to be 0.66eV, which is just greater than half the band gap enthalpy of 1.2eV within the measurement temperature range.

Typical TFT characteristics before and after plasma hydrogenation are shown in figure 1. Two distinct changes can be seen to have occurred:

i) as is commonly found, the TFT on-characteristic showed a considerable improvement in both the sub-threshold slope (1.65V/dec) and threshold voltage (3.2V),

ii) the gate voltage dependence of the off-current was markedly reduced.

The strong variation of drain current with negative gate bias in the non-hydrogenated sample can be best understood by reference to the I_D-V_D characteristic in this regime. This is shown in figure 2 and a linear variation of I_D is seen over a range of V_D until a superlinear regime occurs, the onset of which occurs at lower values of V_D as V_G increases. This more rapid rise of drain current with drain bias occurs only with negative gate bias and previous examination of this effect [3] has ascribed it to field enhanced generation at the induced p-n junction at the drain. Negative gate bias will produce a hole rich region at the surface of the TFT and the linear I_D-V_D characteristic suggests an ohmic hole current between source and drain. This is, at first sight, surprising since the device is contacted by non-injecting n^+ regions; however, if the lifetime within the material is sufficiently low, then the hole generation rate can be sufficiently large to sustain an ohmic hole current in high resistivity material. A hole current with negative gate bias has further been confirmed by observing the change of sign in the Hall coefficient in a Hall TFT as the gate was swept from positive to negative bias. The condition for supporting an ohmic hole current in the device, is that the drift current can be maintained by a generation current. Taking initially the case when the material is intrinsic (assumed to be at the minimum conductivity point) then the hole drift current is

$$J_{dh} = q\mu p_i \frac{V_D}{L} \qquad (1)$$

(assuming $\mu_n \simeq \mu_p = \mu$)

and the generation current is

$$J_g = q \frac{p_i}{\tau} \Delta x \qquad (2)$$

(Δx is width of generating region and τ is the the generation lifetime)

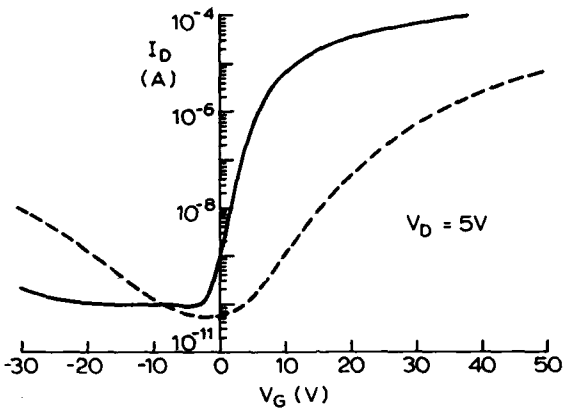


FIGURE 1

TFT transfer characteristics before and after plasma hydrogenation. (-hydrogenated, —unhydrogenated).

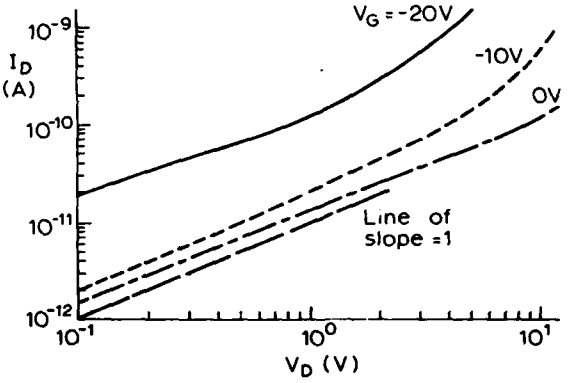


FIGURE 2

I_D-V_D characteristics of the TFT in the off-state for a non-hydrogenated device.

and the maximum generation current will be when $\Delta x = L$. Therefore the condition for the current to be drift current limited and not generation rate limited (as in a normal p-n junction) is that

$$q\mu p_i \frac{V_D}{L} < q \frac{p_i}{\tau} L \qquad (3)$$

$$\text{i.e.} \qquad \tau < \frac{L^2}{\mu V_D} \qquad (4)$$

or $\tau < t_{tr}$ where t_{tr} is the carrier transit time.

An upper limit to the lifetime can be estimated by assuming that the minimum current in the I_D-V_G curve occurs when the layer is intrinsic up to the surface. In that case, if we assume that the free carrier density is n_i+p_i and that $\mu_n \sim \mu_p$, then from equation (1) a value of 32 cm^2/Vs is obtained for the effective mobility of the intrinsic carriers. Substituting this into equation (4) with V_D *of 10V gives* $\tau < 3ns$.

The argument can be extended from the intrinsic case to one in which a hole inversion layer is formed at the surface and equation (4) changes to

$$\tau < \frac{p_i}{\bar{p}} t_{tr} \qquad \left(< \frac{p_i}{\bar{p}} \frac{L^2}{\mu V_D} \right) \qquad (5)$$

where $\bar{p}$ is the mean hole concentration in the inversion layer. Equation (5) is more demanding upon lifetime than equation (4) and it is clear that there must ultimately be a value of $\bar{p}$ for which the inequality is no longer satisfied and the current becomes generation rate limited. Indeed, in figure 2 there is a range of V_D over which the slope of the I_D-V_D curve drops below unity for $V_G = -20V$, indicating the incipient departure from a drift limited current.

The I_D-V_D curves after plasma hydrogenation are shown in figure 3 for $V_G = -5V$ and $V_G = -20V$. For $V_G = -5V$, a linear I_D-V_D curve was found up to $V_D \lesssim 2V$, but unambiguous saturation of the current was seen thereafter, until the previously referenced breakdown occurred. The lifetime estimated from this curve is $\sim 1.3\times10^{-8}s$. The effect is even more clearly seen with the $V_G = -20V$ curve in which current saturation at a level equal to that for $V_G = -5V$ was found for $V_D \gtrsim 0.5V$. What has happened is that the generation lifetime has been increased by the plasma hydrogenation so that the breakdown of the inequality in equations (4) or (5) is seen at reduced values of V_D; and as $\bar{p}$ is increased (by increasing V_G to -20V) so the drain bias at which the hole current saturates at its generation rate limit is correspondingly reduced. In this generation rate limited regime, the TFT off-

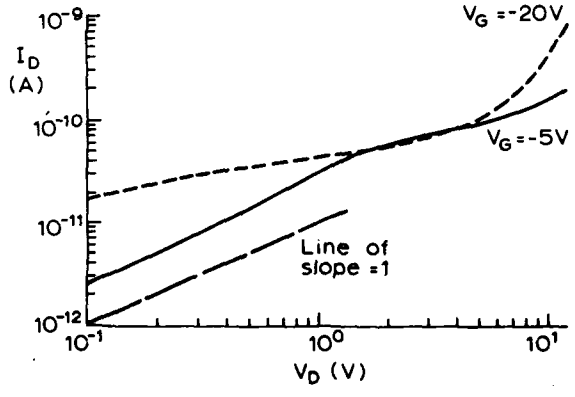


FIGURE 3

I_D-V_D characteristics of the TFT in the off-state after plasma hydrogenation.

current will be determined by the generation lifetime, and increasing this value will reduce the leakage current. Interestingly, for the sample shown here, there is actually a slight increase in the minimum current and this is probably due to an increase in carrier mobility due to a reduction in grain boundary scattering. However, other differently

prepared devices showing a greater change in lifetime have correspondingly displayed a reduction in minimum conductivity.

One consequence of an increase in lifetime within the material is an expected increase in the optical sensitivity of the device and this is shown in figure 4. This is a plot of the photocurrent as a function of white light illumination intensity for devices with and without plasma hydrogenation. The photocurrents are 3-4 times higher in the former device indicating a lifetime increase of comparable magnitude.

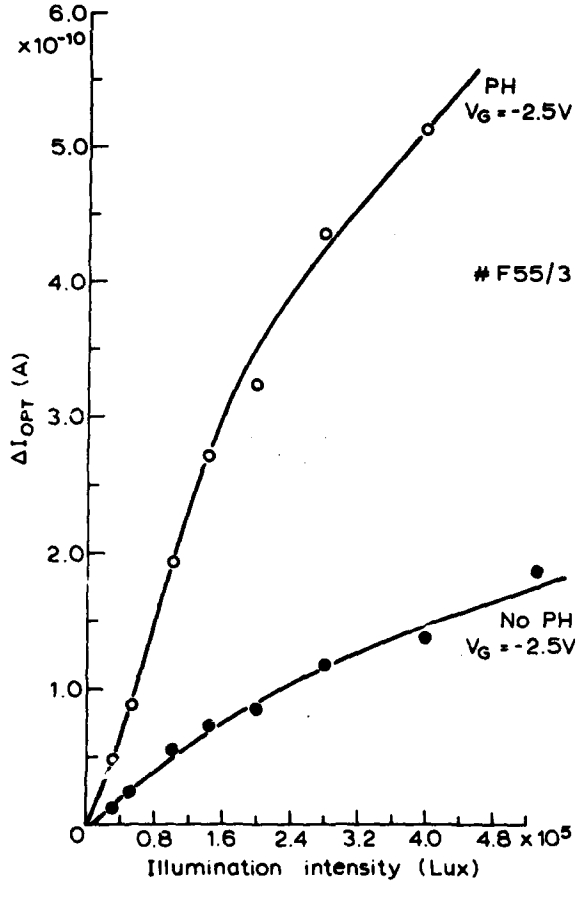


FIGURE 4

Comparison of TFT photocurrent in the off-state with illumination intensity for hydrogenated and non-hydrogenated devices.

4. CONCLUSION

Plasma hydrogenation has been shown to increase the generation and recombination lifetimes of carriers in low temperature processed poly-Si TFTs. In non-hydrogenated TFTs the generation lifetime was sufficiently low to lead to ohmic hole current flow between non-injecting contacts in n-channel devices. The effect of the plasma in increasing the lifetime was to cause the leakage current to become generation rate limited and thereby reduce the influence of the gate potential on the hole leakage current. The increase in carrier lifetime was further confirmed by the concomitant increase in the optical sensitivity of the leakage current.

REFERENCES

[1] Malhi, S.D.S. et al., IEEE Trans. ED-32, 258, (1985).

[2] Aruga, S., Araki, R., Kamakura, H., Shinozaki, J., Morozumi, S., SID'87 Digest, p.75, (1987).

[3] Greve, D.W., Potyraj, P.A., Guzman, A.M., Solid State Electronics, 28, 1255, (1985).

DEVELOPMENT OF CMOS COMPATIBLE P-CHANNEL JUNCTION-FIELD-EFFECT TRANSISTORS FOR VERY LOW NOISE APPLICATIONS

D. Stuch, H. Skapa, H. Vogt and G. Zimmer

Fraunhofer-Institute for Microelectronic Circuits and Systems
D-4100 Duisburg, FRG

CMOS compatible Junction-Field-Effect-Transistors (JFETs) for very low noise applications with different transistor layouts have been developed and tested. The 1/f-corner frequency is below 1 kHz. An equivalent noise voltage density of 1 nV/√Hz at 100 kHz was achieved. The transistors were applied as discrete elements in the inputstage of a transimpedance amplifier for plumbicon tubes. A total signal to noise ratio of 55 dB has been measured.

1. INTRODUCTION

During the last decade VLSI circuits have gained a stage of development towards high reliability, high density, low power consumption and high processing speed. However some applications in the field of analog signal processing additionally require very low 1/f noise, which can hardly be achieved by single MOS transistors. Since low noise specifications can be performed much easier by JFETs than by MOSFETs [2], it is reasonable just to replace the MOSFETs by JFETs in VLSI applications where low 1/f noise is important. In the following sections development, testing and application of CMOS compatible JFETs with low noise characteristics are presented.

2. DEVELOPMENT OF CMOS COMPATIBLE JFETS

2.1. Technology

Figure 1 represents a cross-section of a JFET. N-well formation, active area processing and gate oxide deposition are due to our standard CMOS process [1]. P-channel implantation with boron requires one additional photomask. The implantation energy (150 KeV) and dose (Boron $1 \cdot 10^{12} cm^{-2} - 5 \cdot 10^{12} cm^{-2}$) control

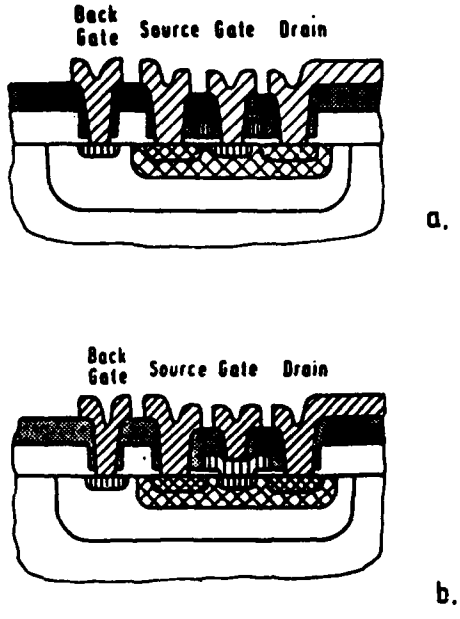


Figure 1 Cross-section of a JFET with two different process sequences.
a.) using PtSi/TiN as ohmic contact to the metal layer
b.) using phosphorous doped polysilicon as ohmic contact to the metal layer

the threshold voltage U_T, maximum drain current I_{DMax} and maximum transconductance g_{mMax}. Topgate and source-drain regions are patterned by two different ways in Fig. 1a. and Fig. 1b. [2, 3]. In figure 1a. the topgate area is defined by one additional photomask and opened via gate oxide etching. Afterwards polysilicon is deposited and doped with phosphorus. The n^+ junction of the topgate is formed by outdiffusion of phosphorus through the topgate area into the p-channel. The polysilicon is patterned by a photomask which is also required for the MOS-transistors. The source-drain regions of the JFETs and of the PMOS-transistors are doped simultaneously. In Fig. 1b. topgate and source-drain regions are patterned by formation of polysilicon spacers. The topgate area of the JFETs and the source-drain regions of the NMOS-transistors are simultaneously doped. In order to reduce parasitic capacitances between the topgate and the channel regions, the polyspacers are removed after the implantations. Since the junction depth of the source drain regions is rather flat (0,3 μ) particular attention has been paid to the metallization level. In order to avoid aluminium spiking and to get a low contact resistance between the junctions and the metallization a triple metal layer PtSi/TiN/Al has been used. Details of the metallization layer are reported in [5].

2.2. JFET layouts

Figure 2 represents two JFET layouts which have been realized. Both layouts are composed of several subtransistors which are connected parallel in the metallization layer. While the source-drain regions are arranged like squares on a chessboard in Fig. 2a they have been designed as parallel stripes in Fig. 2b. Although the overall W/L ratio is about 2000/1.5 in both transistor layouts, the active area of the chessboard type JFET is reduced by 30 % with respect to the active area of the stripe type JFET layout. Therefore parasitic capacitances resulting from backgate junctions are also reduced by the same amount.

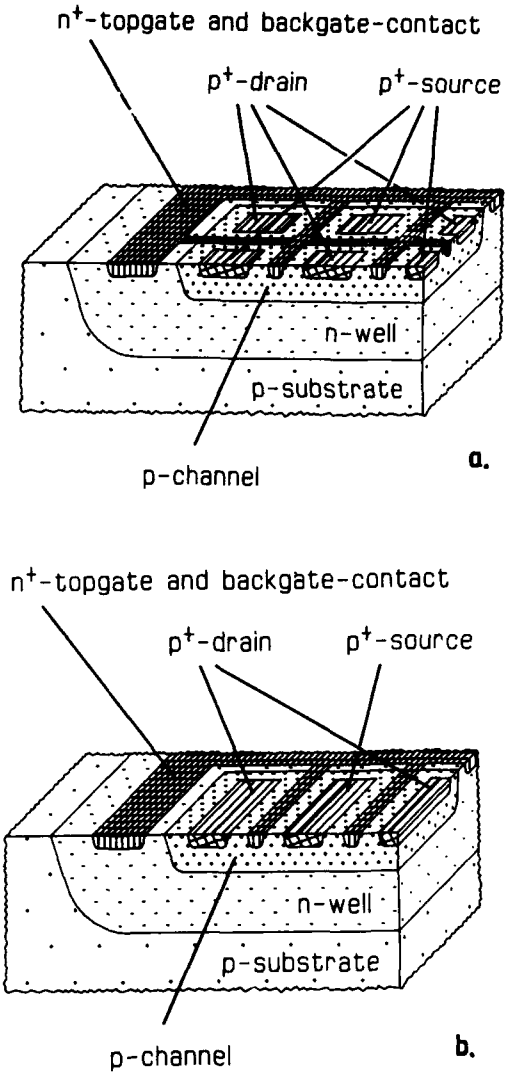


Figure 2 Characteristic JFET layouts
a.) Chess board-type JFET
b.) Stripe-type JFET

3. RESULTS

Figure 3 displays the output-characteristics of a stripe type JFET. Fig. 3a shows that a maximum drain current of 30 mA is reached at zero gate-source bias. The transconductance amounts 22 mA/V at this point. As can be seen from Fig. 3b. the channel is pinched off at 2.8 V. The output characteristics of the chessboard type JFETs is very similar.

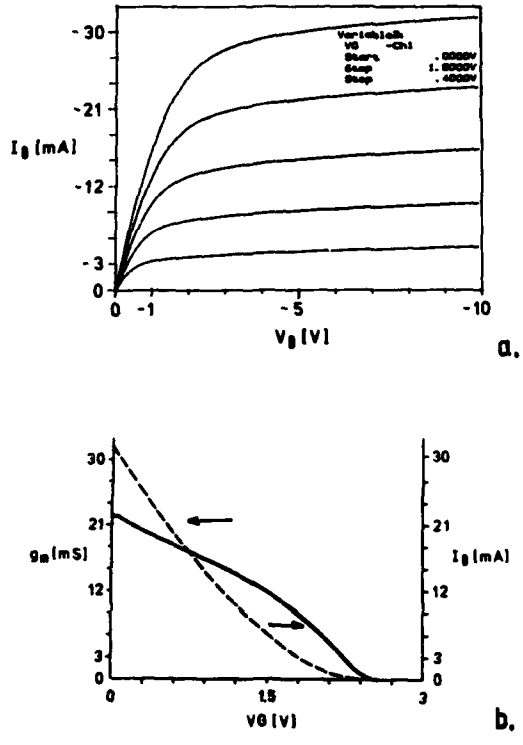


Figure 3 Output-characteristics of a stripe-type JFET.
a.) Drain-source current versus drain-source voltage
b.) Transconductance (——) and drain current (- - -) versus gate-source voltage.

Measurements of the equivalent noise voltage density have been performed using a HP 3583A spectrum analyzer with a low noise cascode preamplifier (Fig. 4). For the JFET under test the operational point was fixed to I_D=10mA; g_m=14mA/V and V_{DG}=-7V. The noise measurements have been corrected for the noise of the measurement equipment. The corrections were in the order of 10 %. Fig. 5 represents the spectral noise voltage density of a stripe type JFET. The 1/f-corner-frequency is below 1 kHz in Fig. 5. The thermal noise of the stripe type JFET is typically 1 nV/√Hz at 100 kHz.

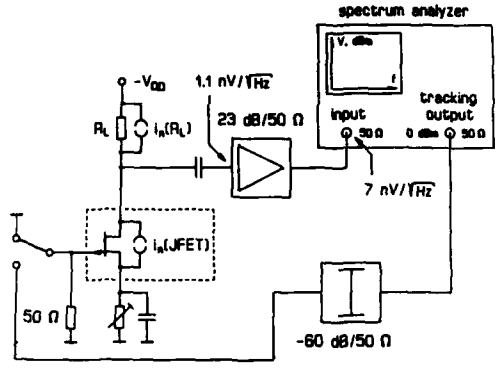


Figure 4 Setup of the noise-measurement-equipment with parasitic noise-sources.

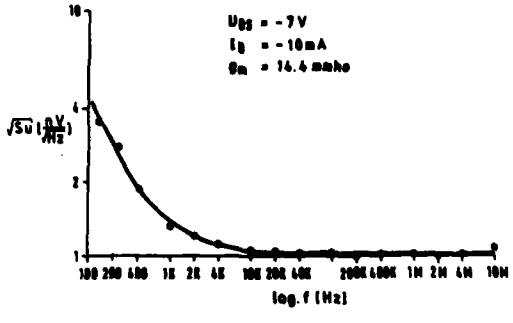


Figure 5 Noise behaviour of a stripe-type JFET

This value agrees rather well with calculations from [5]:

$$\frac{dV}{df} = \sqrt{\frac{8}{3}\frac{KT}{gm}}$$

The thermal noise of the chessboard type JFET is about a factor of 1.4 higher. This behaviour may be referred to the transistor layout in Fig. 2, where the topgate diffusion areas of the subtransistors are organized in a network. Each diffusion area is an ohmic resistance producing thermal noise according to the Nyquist formula

$$\frac{dV}{df} = \sqrt{4KT \cdot R}$$

and transmits it to neighboured subtransistors. In principle the same situtation holds for the stripe type JFET. But there the noise contribution from one subtransistor is only transmitted via a linear line. The overall effect is that in the case of the chessboard type JFETs noise in much more correlated and therefore increased than in the case of the chessboard type JFET.

The stripe type JFETs were applied as discrete elements in the input-stage of a transimpedance amplifier for plumbicon tubes (Fig. 6). The overall signal to noise ratio in this application is 55 dB. This value is limited by the gate-source capacitance which is about 10 pF. In order to reach a signal to noise ratio of 60 dB, this capacitance has to be reduced by at least a factor of two. This can be done by lowering the doping concentration in the channel region.

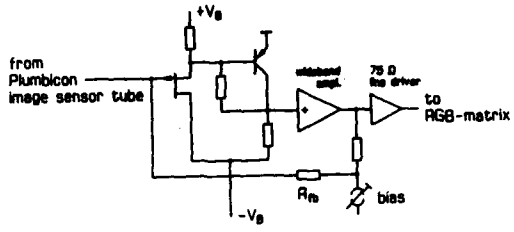


Figure 6 Application of a JFET with low noise characteristics in a preamplifier for plumbicon tubes.

4. CONCLUSIONS

Development of p-channel JFETs was presented. Different transistor-layouts have been realized and testet. JFETs with a stripe type structure provide the best results with respect to noise characteristics. Even better results are expected from n-channel JFETs.

[1] G.Zimmer CMOS-Technologie, Oldenburg 1982

[2] W. Sansen Nuclear Instruments and Methods A253 (1987) 427

[3] H. Vogt Nuclear Instruments and Methods A253 (1987) 434

[4] H. Vogt, G. Zimmer, E. Stein, Essderc 1985

[5] H. Vogt, V. Gruber Essderc 1986

P2.2.6

STATISTICAL MEASUREMENTS OF PMOS SUBTHRESHOLD CURRENT FOR 1.3 TO 0.5 MICRON CHANNEL LENGTHS

T. Ternisien d'Ouville, R. Basset, D.T. Amm, S. Ravazzani, P. Delpech, D. Moi, M. Pacli, B. Minghetti and H. Mingam

Centre National d'Etudes des Télécommunications, Chemin du Vieux Chêne, BP 98, F-38243 Meylan.

The threshold voltage and subthreshold current of micron and submicron PMOS devices have been investigated for both a classical n-well and a retrograde n-well process. The two processes show similar threshold voltage characteristics down to 0.5 micron but the subthreshold current is much improved for the retrograde process.

1. Introduction:

It is well known that in N+ polysilicon gate technology, both N-channel and P-channel FET's require P-type implants in order to produce the proper threshold voltages for VLSI applications. Although PMOS devices are less sensitive to breakdown and hot carrier effects than the NMOS devices, the buried channel PMOS devices are particularly affected by threshold voltage lowering and increased subthreshold current [1].

Various methods have been proposed to prevent this degradation in PMOS performance such as double implantation to create weakly retrograde wells [2], implantation of punchthrough stops [1], or the use of strongly retrograde n-wells [3]. All these methods have the same basis, that is to introduce more charges into the device to slow the expansion of the drain depletion region into the channel region.

This paper presents the threshold voltage and subthreshold current characteristics for a "classical" n-well PMOS process and compares the results with those of a retrograde n-well process.

2. PMOS Process and Devices:

The "classical" n-well process has the maximum n-well concentration at the surface and decreases monotonically into the bulk silicon. A typical phosphorus implantation was 3.3×10^{12} atoms/cm^2 at 180 keV with an anneal temperature of 1200°C. The retrograde n-well process used a high implantation energy of 900 keV (P^{2+} at 450 keV), a dose of 5×10^{12} or 7×10^{12} atoms/cm^2 with a lower anneal temperature (1000 to 1050°C). This yields the peak of the concentration profile at a depth of about 1 micron. Typical dopant profiles are given in Figures 1 and 2 for the source/drain regions, the boron channel implant for threshold adjustment and the n-well concentrations.

All the measurements reported here were made on devices of designed gate lengths of 1.0 to 1.5 microns and widths fixed at 50 microns. During processing, either polysilicon or silicon dioxide spacers were used to lessen the lateral straggling during implantation and diffusion of the source/drain implants into the channel region. The effective gate length could be varied by overetch of the polysilicon gate, or variation of the spacers and the source/drain diffusion. The electrical lengths of the PMOS devices were 0.3 to 0.6 micron less than the designed length on average.

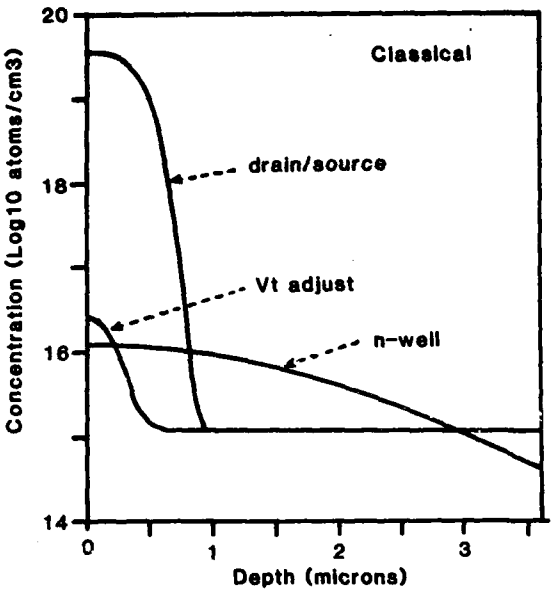


Figure 1
Classical n – well process. Concentration profiles for boron doped drain/source region and threshold voltage implant, and phosphorus doped n – well as calculated by TITAN process simulator.

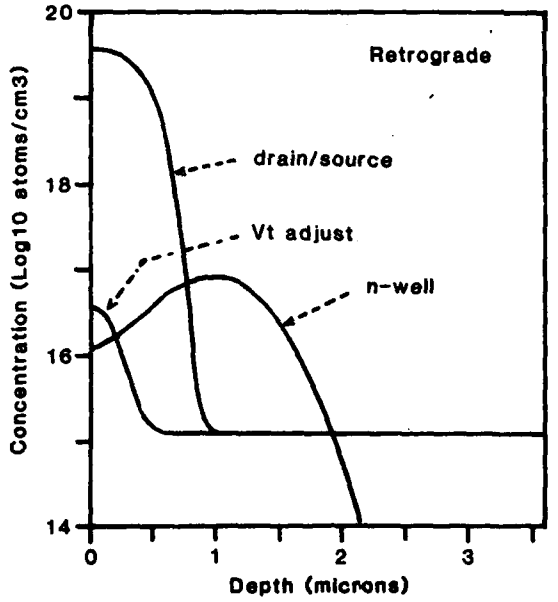


Figure 2
Retrograde n – well process. Concentration profiles for boron doped drain/source region and threshold voltage implant, and phosphorus doped n – well as calculated by TITAN process simulator. (for n – well implant of $7x10^{12}$ and Vt adjust implant of $7x10^{11}$ atoms/cm^2).

3. Results and Discussion:

The threshold voltage was measured for the classical n – well PMOS devices and the results for more than 300 devices are presented in Figure 3 as a function of the electrical length. This figure shows the well – known threshold voltage degradation – as the electrical length is decreased, the device becomes more difficult to turn off. It can be seen from Figure 3 that the threshold voltage is only a weak function of the length above 0.5 micron. It is thus reasonable to assume that this process could produce a controllable threshold voltage for an electrical length of 0.5 micron.

The threshold voltage results for the retrograde n – well PMOS devices are shown in Figure 4. Basically, the retrograde n – well gives the same results as classical n – well down to 0.5 micron. The retrograde n – well has not yet been studied below 0.5 micron and it is not clear whether the two processes would have the same threshold voltage behaviour. Comparing the variation of threshold voltage, it is obvious that the retrograde n – well yields a larger spread in the data. This is probably due to the fact that the surface doping concentration for the retrograde n – well is more sensitive to process variation due to the large concentration gradient near the surface.

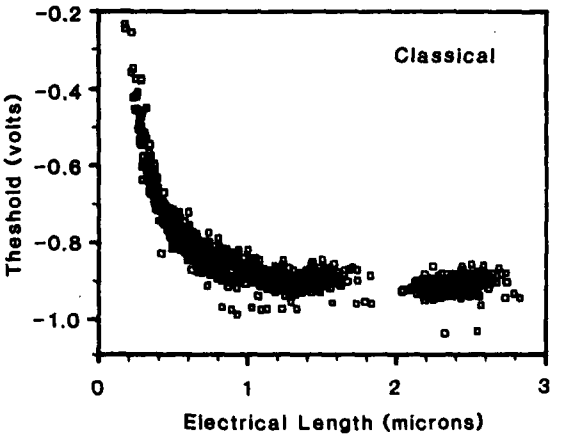


Figure 3
PMOS threshold voltage as a function of electrical length for the classical n – well process.

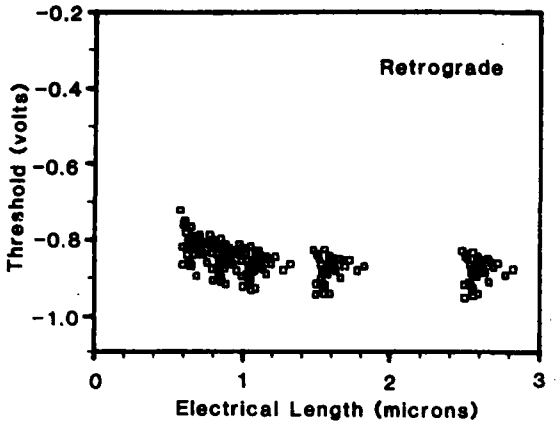


Figure 4
PMOS threshold voltage as a function of electrical length for the retrograde n – well process. (for the profile concentrations given in Figure 1).

The subthreshold current was measured with 0 volts on the gate and a source/drain voltage of – 10 volts. Since the PMOS devices normally performed well, such a large drain/source voltage was needed in order to create currents larger than the automatic test equipment lower limit of $1x10^{-10}$ A. The current characterisitics for the classical and retrograde n – wells are given in Figures 5 and 6 respectively.

The curve in Figure 5 can be characterized by three regions. For lengths greater than 1.1 micron, the current is limited by the measurement system. From 0.8 to 1.1 micron, the current varies exponentially with electrical length. For lengths less than 0.8 micron the current appears to saturate.

The subthreshold current has also been simulated with the TITAN/JUPIN 2D simulator [4,5]. The simulation results are shown as the solid line in Figure 5. These simulations have shown that, as expected, the subthreshold current path is about 0.2 micron into the bulk silicon. Also, the simulations and experience indicate that the exponential dependence of current on length is the standard drain induced barrier lowering (DIBL). The region of current saturation is in fact a resistive region were the barrier has been completely destroyed by the drain voltage.

Choosing an arbitrary minimum acceptable current of $1x10^{-9}$ A (20 pA per micron width) limits device lengths to about 1 micron for the classical well. Although the effects of threshold voltage and subthreshold current are not completely separable, it appears that subthreshold current is a greater problem that the threshold voltage which could be controlled to about 0.5 micron.

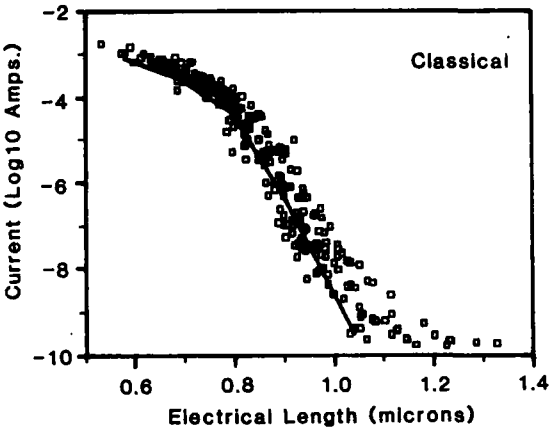


Figure 5
Subthreshold current (Vds = – 10 volt and Vgate = 0 volt) for classical n – well process. The solid line represents the results from 2D simulator JUPIN.

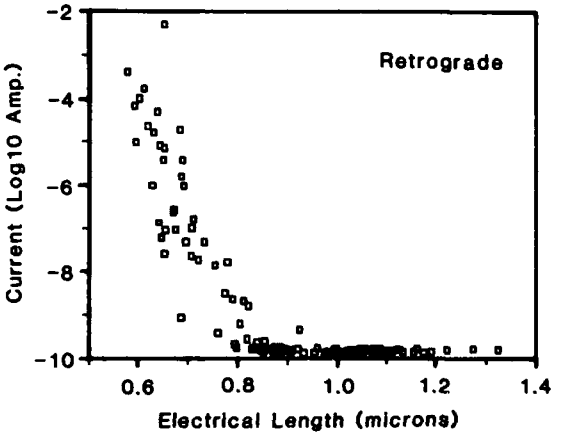


Figure 6
Subthreshold current (Vds = – 10 volt and Vgate = 0 volt) for retrograde n – well process.

Table 1.

Process	Dose Boron ($atoms/cm^2$)	Dose Phosphorus ($atoms/cm^2$)	Vt (at L = 1u) (volts)	Slope (decade/micron)	L_{elec} at $1x10^{-9}$ Amp. (microns)
classic (poly spacer)	$5.5x10^{11}$	$3.3x10^{12}$	−0.87	−21 ± 1	1.05 ± 0.05
classic (SiO_2 spacer)	$5.5x10^{11}$	$3.3x10^{12}$	−0.86	−18 ± 1	1.07 ± 0.07
retrograde	$4.0x10^{11}$	$5.0x10^{12}$	−0.84	−23 ± 2	0.85 ± 0.06
retrograde	$7.0x10^{11}$	$7.0x10^{12}$	−0.87	−20 ± 2	0.82 ± 0.10

The subthreshold current measurements for the retrograde n−well in Figure 6 show only the DIBL region and no saturation or resistive region. The resistive region probably occurs for gate lengths less than those studied here. The DIBL region of Figures 5 and 6 are important in describing how the drain depletion region extends into the channel. This region has been described for various processing conditions and the results are presented in Table 1. For both retrograde and classical n−well processes, the current at Vds = −10 volts increases by about 2 decades per 0.1 micron decrease in gate length. However there is a significant shift in the subthreshold characteristics. On average, the use of the retrograde n−well permits a decrease of 0.2 microns in gate length to give the same current.

The simulation results for the two processes showed that in both cases the subthreshold current is very sensitive to the n−well doping concentration. For example, a 10% decrease in the classical well concentration would cause the current to increase by about one decade. For the retrograde n−well, it is not yet determined whether the concentration gradient near the surface has any effect on device performance or whether the improvements observed are due entirely to the increased doping concentration. Although retrograde n−wells have be promoted mainly for improving latchup immunity and reducing the extent of the well regions [3], it is seen in this work that the retrograde n−well has a beneficial effect on subthreshold current characteristics as well.

4. Conclusions:

For the PMOS processes presented here, the reduction of dimensions is limited by subthreshold current and not so much by the degradation of threshold voltage. Retrograde n−well devices offer a gain of 0.2 microns in gate length or for the same gate length, almost 4 orders of magnitude decrease in subthreshold current at high values of drain voltage.

References:

[1] Odanaka, S. et al, IEEE Elec. Dev., ED−33 (1986) 317.
[2] Cham, K.M. and Chiang, S−Y., IEEE Elec. Dev., ED−31 (1984) 964.
[3] Lai, F−S.J. et al, IEEE Elec. Dev., ED−33 (1986) 1308
[4] Belhaddad, K., Corbex, C. and Poncet, A., NASECODE IV (1985) 193.
[5] Gerodolle, A., Martin, S. and Marrocco, A. NASECODE IV (1985) 287.

INVESTIGATION OF THE Al-ULTRATHIN SiO_2-Si SYSTEM BY COMPARISON OF THEORETICAL AND EXPERIMENTAL CURRENT-VOLTAGE CHARACTERISTICS

Bogdan MAJKUSIAK, Andrzej JAKUBOWSKI and Alfred ŚWIT

Institute of Electron Technology,
Technical University of Warsaw,
Koszykowa 75, 00-662 Warsaw, Poland

Parameters of the Al-SiO_2-Si(n) system with oxide thickness in the range from 2.5 nm to 4.5 nm are determined by fitting theoretical to experimental current-voltage characteristics. The Si-SiO_2 barrier height decreases while the electron effective mass in SiO_2 and the oxide effective charge increase with decreasing the oxide thickness.

1. INTRODUCTION

There is a great interest in properties of the metal-ultrathin SiO_2-Si system. Firstly, it follows from the fact that this system is the basic element of MOS type devices employing the tunneling effect as MOS switching devices, solar cells, bipolar transistors with tunnel emitter. Secondly, the flow of too large tunnel current between the gate electrode and semiconductor substrate limits the trend existing in the MOS/VLSI technology to decrease the gate-oxide thickness. This limitation is especially important for application of MOS structures in DRAM cells because the correct transistor action was observed in transistor with gate oxide of thickness even 2.5 nm [1].

Because of an existance of the transition layer at the oxide-silicon interface with composition and properties different from the bulk SiO_2, parameters of the MOS system with oxide layer several times thicker than this transition layer are expected to depend on the total oxide thickness. Literature data concerning the Si-SiO_2 barrier height X_c generally confirm an existance of such e dependence but scattering of experimentally obtained results is large. According to the internal photoemission measurements the barrier height in the case of ultrathin oxide of thickness 4.3 nm is the same as for "thick" oxides [2]. On the other hand, experimental works based on analysis of tunnel current suggest that X_c decreases with decreasing the oxide thickness [e.g.3-5].

It is the aim of this work to draw conclusions on the dependence of some physical parameters of the Al-SiO_2-Si system on oxide thickness from fitting the current-voltage characteristics theoretically determined to the experimental ones.

2. EXPERIMENTAL

The MOS diodes were fabricated simultaneously except the ultrathin oxidation process. Phosphorus-doped and (100) oriented silicon wafers of 2.7 Ωcm resistivity after chemical cleaning and final etching in dilute HF were steam oxidized at 1000 °C to form the field oxide layer. Then the windows of 390 μm diameter were opened in the field oxide by photolithography. The wafers were oxidized separately (but together with control wafers) in dry O_2 at 800 °C with 5-min preheat and 10-min final annealing in argon ambient. The oxidation time was varied (in sequence: 22,15,9,4 and 1 min) to give various oxide thicknesses. Immediately after oxidation the oxide thickness was measured ellipsometrically on the control wafer (results: 4.5, 3.9, 3.4, 2.9 and 2.5 nm) and the wafer for further processing was loaded to the vacuum system. After compliting the whole

series of oxidation processes the Al film was resistively evaporated simultaneously on all wafers. Then, the Al electrodes were formed by photolithography, the oxide was removed from the back side of the wafers and Al back contacts were evaporated.

The I-V characteristics of the diodes were measured without any electrical forming. They were stable and reproducible on the wafer provided the diodes were good from the beginning of the measurement.

3. SHORT CHARACTERIZATION OF THE MODEL

The model used to produce the theoretical I-V characteristics is a result of compromise between simplicity and accuracy of the description [6]. In order to avoid numerical solving of the transport equations in the semiconductor region the description of it is based on the assumption of constant position of quasi-Fermi levels across the semiconductor space-charge region. The Seiwatz-Green's description of this region with the deviation of the hole quasi-Fermi level from the bulk Fermi level as an additional parameter is used. The densities of tunnel currents between the metal and allowed energy bands in semiconductor are calculated from the one-dimensional integral formula. For example, the electron tunnel current is given by:

$$J_{CT} = q \int_0^{E_{x\,max}} P(E_x)\, N(E_x)\, dE_x$$

where E_x is the electron energy component associated with the motion normal to the barrier plane, $N(E_x)$ is the "supply function" defined in [7] and the tunneling probability $P(E_x)$ is expressed by WKB exponential factor using two-band barrier model with one average effective mass m_i:

$$P = \exp\left\{-2\left(\frac{2m_i}{\hbar^2}\right)^{1/2} \frac{E_{gi}}{16qF_i}\left[\sin 4\alpha_2 - \sin 4\alpha_1 - 4\alpha_2 + 4\alpha_1\right]\right\}$$

Here E_{gi} is the insulator band gap, F_i is the electric field in the insulator region and $\alpha_{1,2} = \arccos(\phi_{1,2}/E_{gi})^{1/2}$. $\phi_{1,2}$ are the barrier heights $E_{ci}-E_x$ at the turning points x_1 and x_2, respectively. The image forces are neglected as it is suggested in [8]. Therefore, the real potential barrier is modelled by a trapezoidal barrier.

The tunnel current between the metal and interface traps is calculated by adding the contributions from discrete traps levels E_{it} of arbitrarily assumed density $N_{it}(E_{it})$ [9]:

$$j_{ST}(E_{it}) = q\, N_{it}\, (f_t - f_m)/\tau_T$$

where τ_T is the tunneling time constant.

4. RESULTS

Fig.1 shows the measurement points of characteristics representative for respective oxide thicknesses. The total tunnel current in the Al-SiO_2-Si(n) system is dominated in the whole bias range by current of electrons tunneling between the metal and silicon conduction band. The number of parameters affecting this current is large, especially for the reverse bias direction, where the potential distribution in the structure can be additionally affected by non-equilibrium conditions for minority carriers. For oxide thicknesses $t_i > 3.4$ nm there is a voltage range where the slope of reverse characteristics increases what means that the inversion layer forms in the semiconductor surface region. For $t_i < 3.4$ nm the thermal generation processes in the semiconductor are not effective enough to maintain this inversion layer and the surface region is still in deep depletion state due to the tunnel outflow of holes to the metal electrode. To avoid considering the balance between the fluxes of minority carriers and to limit in this way the number of parameters, fitting was performed for positive bias direction, where the semiconductor surface region is in the accumulation state. Additionally under these conditions the interface traps charge variable with voltage is much smaller

than the accumulation layer charge and does not influence electrostatically the potential distribution in the structure and hence, the electron tunnel current. Similarly, the I-V characteristic of the structure in the accumulation state does not depend also on the value of impurity concentration assumed in calculations. It is easy to prove that the potential barrier shape and position of the electron Fermi level at the interface, which together determine the current of electrons tunneling from the semiconductor conduction band to the metal, do not change with the impurity concentration at a given gate voltage in the bias range of strong accumulation.

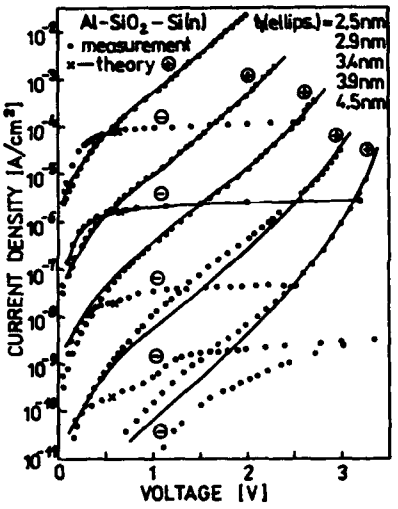


FIGURE 1

The experimental and theoretical current-voltage characteristics.

Fig.1 shows the results of theoretical calculations (solid lines). It was not possible to fit all characteristics by changing only oxide thickness. The characteristics for t_i=4.5 nm and 3.9 nm were obtained after assuming the Si-SiO_2 barrier height X_c=3.05 eV, the Al-SiO_2 barrier height ϕ_{mi}=3.2 eV, the insulator band gap E_{gi}=8.64 eV [10] and the average two-band effective mass in the SiO_2 m_i=0.39 m_0 for t_i=4.5 nm and 0.412 m_0 for t_i=3.9 nm. It was not possible to obtain very good agreement of shape of characteristics for these two oxide thicknesses even when the image forces and different electron and hole effective masses in SiO_2 were taken into account. In order to obtain the required shape of characteristics for diodes with oxide of thickness below 3.9 nm the barrier heights had to be reduced in comparison with values typical for "thick" oxides. We assumed simultaneous reduction of all barrier heights so as to maintain constant value of the metal-semiconductor work function difference. To obtain measured levels of current, the effective mass m_i had to be increased with decreasing the oxide thickness. In order to estimate the oxide effective charge Q_{eff}, fitting was also done for the points of characteristics (marked by crosslets) for which the surface potential is equal to the Fermi potential. The oxide effective charge also had to be increased with decreasing the oxide thickness. Fig.1 shows also a result of exemplary fitting to the characteristic of the diode with oxide t_i=2.9 nm in the whole bias range, for which the interface traps density $D_{it}=4\cdot10^{11}$ $cm^{-2}eV^{-1}$ had to be assumed.

Fig.2 shows the values of X_c, m_i and Q_{eff} obtained from fitting as a function of oxide thickness.

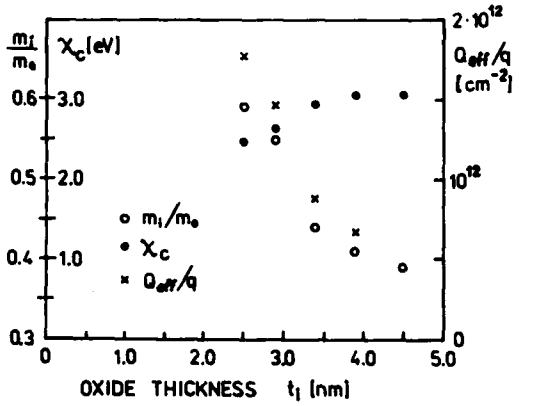


FIGURE 2

Values of the Si-SiO_2 barrier height X_c, electron effective mass m_i and oxide effective charge Q_{eff} resulting from fitting versus the oxide thickness t_i.

5. CONCLUSIONS

The dependence of the Si-SiO_2 barrier height and of the oxide effective charge is rather in agreement with expectactions and literature results. The value of X_c for $t_i \geq 3.9$ nm agrees with result of photoemission measurement [2]. The behaviour of the electron effective mass in SiO_2 with the oxide thickness is the most discussive result. The value of m_i is smaller than 0.6365 m_0 obtained in [10]. However, values of m_i reported in the literature concern tunneling through oxide regions very close to the SiO_2-Si or the metal-SiO_2 interface because they are determined from analysis of the Fowler-Nordheim tunnel current. The greater the electric field in the oxide is, the less the influence of oxide region beyond the transition region on the tunneling probability is. Our results indicate that m_i increases with decreasing t_i what can be a result of increasing part of the transition SiO_x region in the total oxide layer. It is interesting to note that the longitudinal electron effective mass in Si (100) m_l=0.917 m_0.

REFERENCES

[1] Horiguchi S., Kobayashi T., Miyake M., Oda M. and Kiuchi K., IEDM Techn. Dig., p.761 (1985).

[2] Dressendorfer P.V. and Barker R.C., Appl. Phys. Lett. 36, 933 (1980).

[3] Lewicki G. and Maserjian J., J.Appl.Phys. 46, 3032 (1975).

[4] Card H.C., Solid State Commun. 14, 1011 (1974).

[5] Kasprzak L.A., Laibowitz R.B. and Ohring M., J.Appl.Phys. 48, 4281 (1977).

[6] Majkusiak B, Ph.D.Dissertation, Technical University of Warsaw, May 1985.

[7] Majkusiak B. and Jakubowski A., J. Appl. Phys. 58, 3141 (1985).

[8] Weinberg Z.A. and Hartstein A., Solid State Commun. 20, 179 (1976).

[9] Freeman L.B. and Dahlke W.E., Solid-St. Electron. 13, 1483 (1970).

[10] Av-Ron M., Shatzkes M., DiStefano T.H. and Gdula R.A., J.Appl.Phys. 52, 2897 (1981).

INTERFACE STATE ANALYSIS OF MOSFETS WITH A MODIFIED CHARGE-PUMPING TECHNIQUE

G. Przyrembel, W. Krautschneider*), W. Soppa and H.-G. Wagemann

Technische Universität Berlin
Institut für Werkstoffe der Elektrotechnik
Jebensstr. 1, D-1000 Berlin 12, West-Germany
*)now with Siemens Corporate Research Laboratories
Otto-Hahn-Ring 6, D-8000 Munich 83

In this paper we present a modified charge pumping technique to obtain the energy dependence of interface states of MOSFETs. The steps of gate potential shift periodically between 3 voltage levels.

The duration of the gate-mid-level V_{GMID} gives a precise control of the electron and hole emission times avoiding to generate and measure transition times of trapezoidal pulses. Furthermore an extension of charge pumping measurements on chains of MOSFETs in parallel is realized. CV measurements at varactors of the same testchip are evaluated and compared to the results of the charge pumping technique.

1. INTRODUCTION

Due to the reduction of MOSFET dimensions towards smaller geometries, the investigation of the Si-SiO_2 interface becomes important again. This is the consequence of new technology processes and of working conditions under higher electric fields, which both can lead to additional interface states. For monitoring these effects an appropriate measurement method is required.

Charge pumping offers the possibility to obtain the energy dependence of interface states even for submicron MOSFETs. This analysis, proposed by Groeseneken et al. [1], uses the variation of the rise and fall times to control the electron emission time t_e or the hole emission time t_h respectively for achieving a scan of the density of interface states D_{it} over the bandgap.

The electron and hole emission times t_e and t_h are time intervals in which the surface states are discharged and charged respectively, described by the model after Shockley, Read and Hall [2]. On this model the charge pumping effect is based. But hitherto certain discrepancies in the energy distribution of interface states obtained by different methods of analysis still remain. The reason for inaccuracies within the results from charge pumping is connected to the intermediate state of inequilibrium at the semiconductor surface. These time intervals have to be precisely controlled.

2. MODIFIED CHARGE PUMPING TECHNIQUE

An accurate method to control the electron emission time t_e or the hole emission time t_h can be realized by a 3-level pulse train depicted in fig. 1.

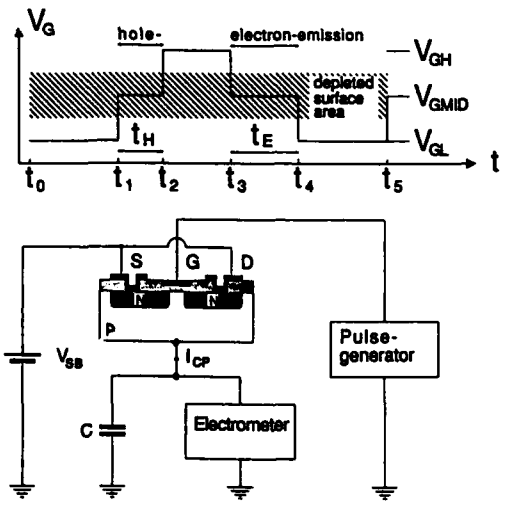


FIGURE 1

3-level charge pumping, diagram of gate voltage pulsetrain and measurement setup.

The purpose of each of the three voltage levels during one period of charge pumping will be described next for the case of a n-channel MOSFET.

1. s t a r t i n g p h a s e V_{GL}
   $0 \leq t \leq t_1$:

the lowest level accumulates positive charge at the surface of the MOSFET and completes the positive charging of the interface states by h o l e c a p t u r e .

2. t r a n s i t i o n p h a s e V_{GMID}
   $t_1 \leq t \leq t_2$:

the mid-level V_{GMID} has to be adjusted to deplete the surface from free carriers, so that no capture processes can occur.

During the interval t_H of V_{GMID} following the switching from V_{GL} to V_{GMID} there will be only h o l e e m i s s i o n from the interface states.

3. t e r m i n a t i n g p h a s e V_{GH}
   $t_2 \leq t \leq t_3$:

the highest level V_{GH} must ensure strong inversion of the surface, so that negative charging of the traps will now be completed by e l e c t r o n c a p t u r e . V_{GH} must be carefully adjusted to appropriately higher values for still reaching inversion if a reverse voltage V_{SB} between the source and drain electrodes to the substrat will be applied.

For the second part of the pulse train accordingly to the first part of the description 3 phases of charge pumping procedure occure

1. s t a r t i n g p h a s e V_{GH}
   $t_2 \leq t \leq t_3$:

negative charging of the surface by electrons within equilibrium

2. t r a n s i t i o n p h a s e V_{GMID}
   $t_3 \leq t \leq t_4$:

electron emission from the interface states within in-equilibrium

3. t e r m i n a t i n g p h a s e V_{GL}
   $t_4 \leq t \leq t_5$:

hole capture by the interface states to equilibrium.

For easily positioning V_{GMID} within a voltage range in which the surface has to be depleted a additional reverse voltage V_{SB} between the source and drain electrodes referring to the substrate is recommended. By this the range for V_{GMID} can be enlarged to several volts as can be seen in fig. 2.

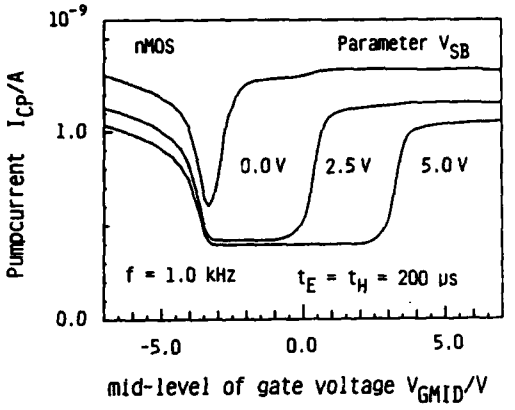


FIGURE 2

Charge pumping current ICP vs. mid-level V_{GMID} of gate voltage at varios source/drain to substrat voltages. The plateaus of minimum current mark the voltage ranges, in which V_{GMID} makes the surface depleted.

When V_{GMID} comes out of the voltage range, in which the surface will be depleted, the fast rise times of the pulses will be responsible for shorter emission times and consequently for a higher pump current. As a contrast to the conventional charge-pumping-technique non-equilibrium processes are observed during phases of constant gate-voltage, not during a ramp period. The time intervals t_E and t_H of the modified pumping method are directly related to the emission times t_e and t_h for electrons and holes respectively, because the capture processes start exactly and abruptly after t_E and t_H, when strong inversion or accumulation are reached. This is expressed in eq.(1).

$$\text{electron emission time } t_e = t_E$$
$$\text{hole emission time } t_h = t_H \qquad (1)$$

By investigating the influence of V_{GL} or V_{GH} on the charge pumping current it is advantageous that the times t_e and t_h are not affected by

varying of V_{GL} or V_{GH} as it will be the case by the conventional trapezoidal pulsetrain. This can be seen by presenting the corresponding equation to eq. (1) for the trapezoidal pulsetrain $t_e = (V_{FB}-V_T)/(V_{GH}-V_{GL}) \times t_f$, [1], where t_f is the falltime from V_{GH} to V_{GL}.

Furthermore, for the modified technique the times t_e and t_h are independent from the voltage interval $V_{FB}-V_T$.

In summary the advantages of the 3-level pulsetrain for controlling the emission times t_e and t_h are

- equality of the pulse interval t_E, t_H and the emission time t_e, t_h, respectively.
- pulsetrain can be generated by two simple generators with fixed rise times; the time intervals t_E, t_H are easier to measure than rise or fall times.
- a very large ratio t_E/t_H is easy to realize.

3. EXTENSION OF CHARGE PUMPING MEASUREMENTS ON CHAINS OF MOSFETS

Efforts were made to enable charge pumping measurements at single devices of MOSFET chains. This is necessary because most of the available testchips only offer chains, in which gates and sources of the MOSFETs are interconnected in parallel. The investigation of a single device is of interest e.g. especially when monitoring the effects of hot carrier stress. To activate a single MOSFET only to charge pumping the following steps must be carried out and this refers to a circuit as shown in fig. 3.

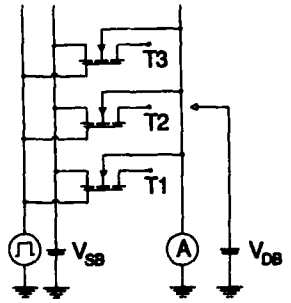


FIGURE 3

Circuit scheme for pumping a single device of a MOSFET chain. T2 is selected.

1) The level of the gate pulse, which provides inversion of the surface, should be chosen some tenth of a volt above threshold. Threshold voltage is given here by assuming V_{DB} as the acting backgate voltage.
2) The common source voltage V_{SB} must be reverse biased high enough, so that no inversion can occur by assuming V_{SB} as the acting backgate voltage. The measured substrate current must approach zero by having all drain contacts floating. Care must be taken, that V_{SB} will not be chosen too high so that the effective channel voltage $V_{DB}-V_{SB}$ causes a substrate current that exceeds the pump current.
3) To select a MOSFET for pumping, the drain electrode must be connected to the V_{DB} voltage. The drain electrodes of the other devices may be left floating or can be connected to the V_{SB} voltage.

It is worth mentioning that the selected device may have a channel current several times higher than the pump current, without affecting the latter.

4. EXPERIMENTAL RESULTS

We have used the new capability of pumping a single device within a MOSFET chain to investigate the dependence of the interface state density $\overline{D_{it}}$ from the effective channel length L_{eff}. In fig. 4 it can be seen, that there is no remarkable change of $\overline{D_{it}}$ to shorter channel length.

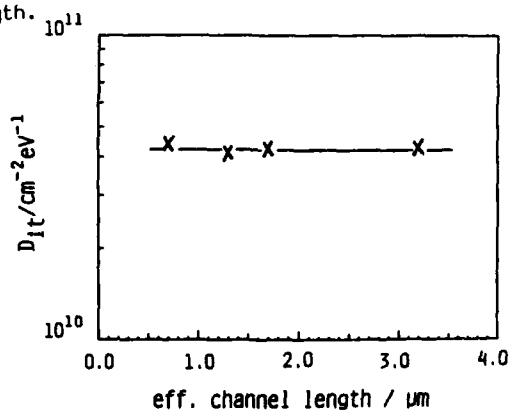


FIGURE 4

Interface state density $\overline{D_{it}}$ as a function of effective channel length L_{eff}.

For applying the modified charge pumping technique to determine the energy distribution of interface states, the times t_E and t_H were varied and the pump current I_{CP} monitored. The results are illustrated in fig. 5.

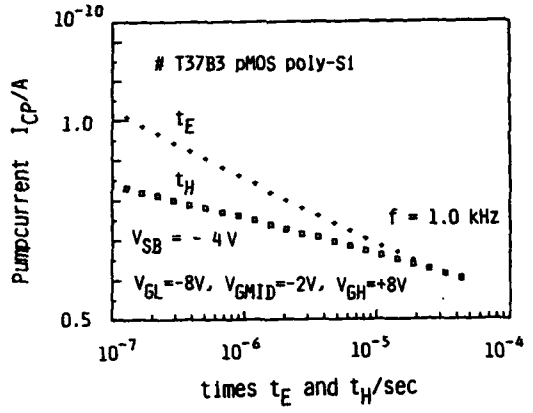


FIGURE 5

Pumpcurrent I_{CP} as a function of the times t_E and t_H, which are defined in fig. 1.

In an analogous way to [1] the slope of the curves is proportional to the density of states at an energy W, which is determined by the values of t_E and t_H, respectively.

A plot of the energy distribution $D_{it}(W)$ of interface states, obtained with the modified charge pumping technique, is shown in fig. 6

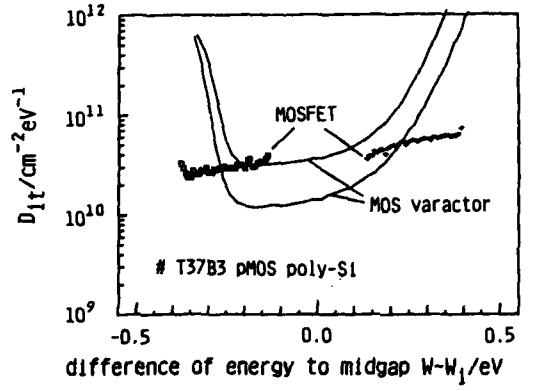


FIGURE 6

Energy distribution of interface states. Comparison between charge pumping method with evaluation after [1] and quasistatic C-V method. MOSFETs and varactors are on the same wafer.

To verify the D_{it} values of a polysilicon gate MOSFET a comparison with results of quasistatic C-V or Q-V [3] measurements of varactors on the same wafer were made. In fig. 6 the results of the $D_{it}(W)$ evaluation from C-V measurements of varactors are also shown. The two curves indicate the spread for different specimens on the same wafer. The increase of $D_{it}(W)$ to the band-edges as obtained from the C-V method has not been observed for the charge pumping results. But the accuracy of the C-V method is restriced to depletion and weak inversion, as is well known. For the investigated varactors this will be the range from -0,29 to +0,29 eV.

5. CONCLUSION

A 3-level-charge pumping technique facilitates the calibration of the non-equilibrium region of band-bending and the evaluation of electron or hole emission times. It provides energy distributions of interface state density comparable to results from C-V-measurements. This technique is applicable to test structures of parallel chains of MOSFETs.

ACKKNOWLEDGEMENTS

For device preparation we appreciate the aid of Dr. D. Haack, Fa. R. Bosch, Reutlingen. Furthermore we thank Dr. W. Seifert and his team for intensive technology assistance. For financial support we are grateful to the Institut für Mikrostrukturtechnik der Fraunhofer-Gesellschaft Berlin/IMT.

REFERENCES

[1] Groeseneken, G., Maes, H.E., Beltran, N., De Keersmaecker, R.F., IEEE Trans ED-31, (1984) p. 42-53

[2] Shockley, W., Read, W.T.Jr., Phys. Rev. 87 (1952) p. 835-842

[3] Ziegler, K., Klausmann, E., Appl. Phys. Lett. 26, (1975), 400

MONITORING THE IMPURITY PROFILE AND THICKNESS OF SEMICONDUCTOR LAYERS WITH THE CHANNEL CONDUCTANCE OF BURIED CHANNEL FIELD EFFECT DEVICES

G.J.L. Ouwerling and M. Kleefstra

Delft University of Technology, Faculty of Electrical Engineering,
Laboratory of Electrical Materials, Mekelweg 4, P.O.Box 5031,
2600 GA DELFT, The Netherlands

A new method to determine the width of depleted layers and the impurity profile in semiconductor materials is presented. The method employs the channel conductance of buried channel field effect devices (with junction, oxide or Schottky-barrier gate isolation) and can be performed with DC measurements only. A concise derivation of the expressions for the impurity concentration and the depletion layer width dependent on the channel conductance is given. Experimental results on both measured and synthetic data are provided.

1. INTRODUCTION

After the processing of semiconductor wafers, it is important to determine the width and the impurity atom concentration profile of epitaxially grown or ion-implanted semiconductor layers.

Probably the most widely used non-destructive method to obtain this information is the differential capacitance or CV-method, as originated by Schottky [1] and elaborated upon by many others, e.g. [2-4]. This method has the advantage of being generally applicable; as a measurement device a reverse biased Schottky- or pn-diode, or a MOS-structure, suffices. However, especially in production environments the following disadvantages must be noted:

- the need to eliminate stray and peripheral capacitances;
- the influence of series resistance (especially in the case of low impurity concentrations);
- the use of an AC test signal which makes the measurement sensitive to interference from other AC sources.

We present a novel profiling method that employs the channel conductance versus the gate voltage of buried channel field effect devices. Such a device might either be a JFET, a MESFET or a buried channel MOSFET, devices all described by Sze [5]. These test structures can

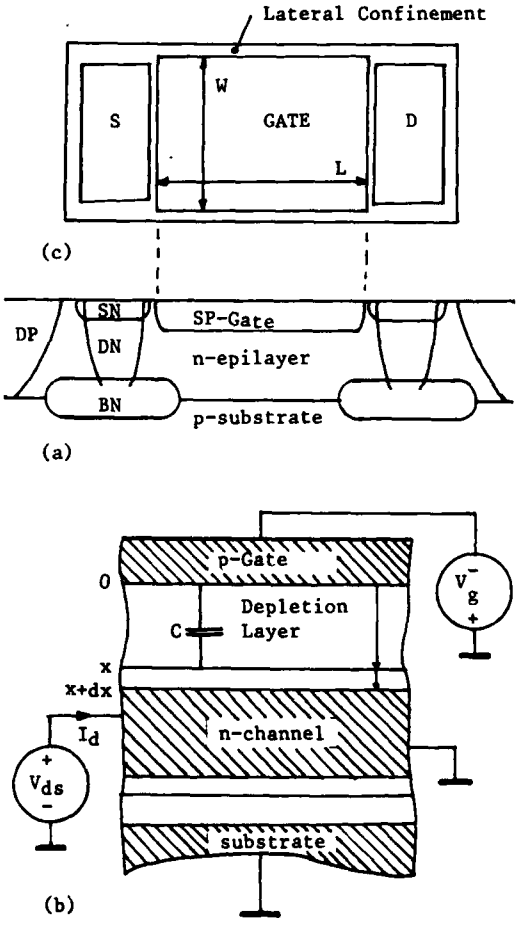


Fig. 1. Channel Conductance Device.
(a) JFET Device in Bipolar process.
(b) Configuration during measurement.
(c) Dimensions of active gate area.

be realized in most bipolar and MOS IC fabrication processes. Therefore, this channel conductance method also covers a large range of possible applications. Moreover, it offers the following advantages over the CV-method:

- rapid DC (IV) measurements only;
- less influenced by side effects, hence smaller devices;
- especially suitable for low impurity concentration measurements.

2. DEVICE STRUCTURE AND PRINCIPLE OF THE METHOD

In this paper we restrict ourselves to the case of a Junction FET measurement device. Our experimental results have also been obtained with such a device. The appropriate expressions for MESFETs are equal to those of JFETs; the derivation of the expressions for MOSFETs is straightforward.

Fig. 1a depicts a cross section through a JFET channel conductance device fabricated in a common bipolar process. The gate is formed by a shallow p-type (base) diffusion. The target of the measurement is the thickness and the impurity profile of the n-type epitaxial layer, which constitutes the channel of the JFET. To eliminate unwanted series resistance, source and drain are formed by shallow n-type (emitter), deep n-type (collector wall) and n-type buried layer diffusions.

A negative voltage V_g is applied to the gate in order to build up a depletion layer in the channel. A small voltage V_{ds} is applied between source and drain in order to measure the channel conductance $G = I_d/V_{ds}$ dependent on V_g. This is schematically depicted in fig. 1b.

In our derivation, the following assumptions are made:

1. The influence of non-zero V_{ds} on the channel conductance is negligible;
2. The extension of the depletion layer in the gate may be neglected;
3. An abrupt space charge edge separates the depleted region from the channel region (the depletion approximation).

A method to account for the (small) influence of V_{ds} is, in a similar context, provided by Buehler [6]. Both assumptions 2. and 3. are also made in the case of the CV method [1]. For the error caused by the depletion approximation, correction schemes are available, see e.g. [2].

The traditional expressions for the impurity concentration N and the depletion width x as computed from the small signal depletion layer capacitance C are

$$N(x) = -\frac{C^3}{q\epsilon A^2 \frac{dC}{dV_g}} \quad , \quad x = \frac{\epsilon A}{C} \tag{1}$$

where A is the area of the active gate and ϵ the permittivity of the material. By definition C also equals

$$C = \frac{dQ}{dV_g} \tag{2}$$

where dQ is the incremental charge present in the slab with thickness dx that is added to the depletion layer when V_g is increased with dV_g,

$$dQ = qAN(x)dx \quad \text{(see fig. 1b).} \tag{3}$$

The same increment dV_g decreases the channel conductance G with an amount

$$dG = q\mu F_g N(x)dx \tag{4}$$

where μ is the low-field mobility of the majority carriers (here electrons). F_g is a dimensionless geometrical factor that is solely determined by the shape of the active gate region. In the case of a rectangular gate, F_g equals the aspect ratio W/L (fig. 1c). From (2), (3) and (4) we find an expression for the capacitance expressed in the conductance

$$C = \frac{A}{\mu F_g}\frac{dG}{dV_g} . \tag{5}$$

Substitution of (5) in the parametric relations (1) yields the desired expressions of N and x in terms of the conductance G:

$$N(x) = -\frac{\left[\frac{dG}{dV}\right]^3_g}{q\varepsilon\mu^2 F_g^2\left[\frac{d^2G}{dV^2_g} - \frac{1}{\mu}\frac{d\mu}{dV_g}\frac{dG}{dV_g}\right]}, \quad x = \frac{\varepsilon\mu F_g}{\frac{dG}{dV_g}} \qquad (6)$$

It may be of interest to note that these formulas can also be derived without the use of the CV expressions (1).

If the low field mobility μ as a function of x (and thus V_g) may be considered constant, e.g. in the case of low impurity concentrations, expressions (6) may be simplified to

$$N(x) = -\frac{\left[\frac{dG}{dV}\right]^3_g}{q\varepsilon\mu^2 F_g^2\frac{d^2G}{dV^2_g}}, \quad x = \frac{\varepsilon\mu F_g}{\frac{dG}{dV_g}}. \qquad (7)$$

These expressions can be evaluated directly.

If μ is not constant, an empirical relationship describing $\mu = f(N)$, an example of which is provided by Masetti [7], must be used in order to find both μ and N using the derived expressions (6). This involves an iterative scheme in which, starting with constant μ, alternatively N and $\mu = f(N)$ are computed until convergence is reached. A further detailed discussion of this scheme is not within the scope of this paper; however, some results are shown below.

3. IMPLEMENTATION OF THE METHOD

Some practical aspects of the new method are discussed below.

If the channel conductance device is fabricated in a process with junction isolation, the measurements will be distorted by the parasitic parallel conductance along the lateral confinement (fig. 1c). Two methods can be thought of to avoid this distortion: compensation by subtraction (fig. 2a) or elimination of the confinement by using a circular (or polygonal) gate (fig. 2b).

Numerical determination of the 2nd derivative of a measured quantity is notably difficult. We used a simple current compensation measurement technique, involving the subtraction of a constant current from I_d, to accurately determine conductance differences ΔG as a function of ΔV_g. Differentiation of this data by curve-fitting (see Ralston, [9]) then provides for a stable and precise calculation of d^2G/dV_g^2.

4. EXPERIMENTAL RESULTS

Results are available on both measured and synthetic channel conductance data. Measurements were made with a channel conductance Process Control Module (PCM), consisting of a circular JFET (fig. 2b) and meant for use in the Junction

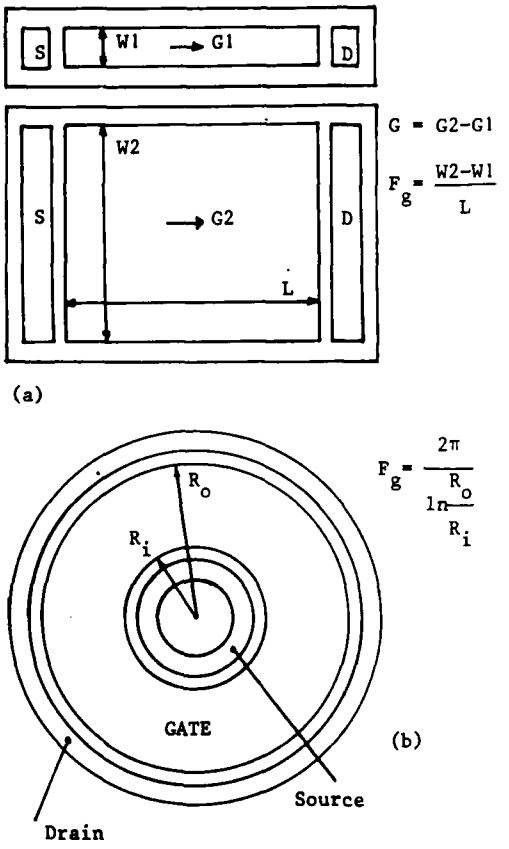


Fig. 2. Compensation of lateral conductance.
(a) With two rectangular devices.
(b) With a circular device.

Charge-Coupled Device bipolar process [10]. The device easily fitted into a standard PCM unit of 500 x 500 μm^2 size. Fig. 3a presents the impurity profile obtained with this module as compared to the profile conventionally obtained with the CV method.

In fig. 3b, a result on synthetic data is shown. Depicted are the impurity profile used to simulate the measurement (a), and the impurity profiles obtained from the simulated conductance data by the above mentioned iterative scheme (b-e). Clearly, after a few iterations the measured profile coincides with the original one within a few precents.

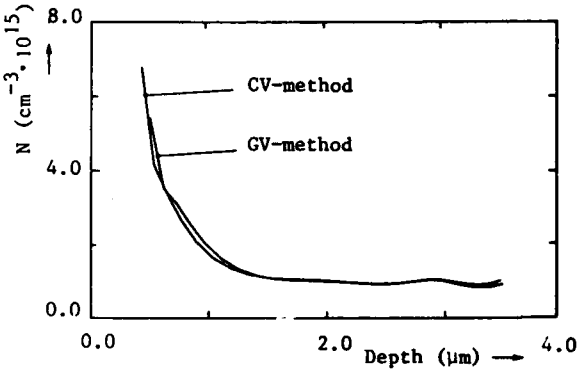


Fig. 3a. Impurity profiles with CV and Channel Conductance (GV) methods.

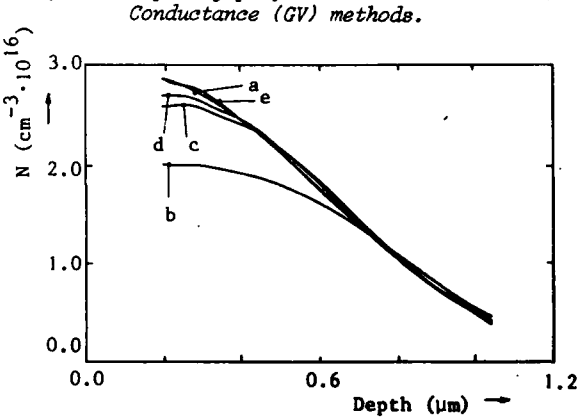


Fig. 3b. Iterative determination of the impurity profile in the case of non-constant mobility.
b - initial guess, c - after 1 iteration, d - 2 iterations, e - 5 iterations. a - desired result.

5. CONCLUSION

Determination of impurity profiles with channel conductance measurements provides an attractive alternative to profiling with capacitance-voltage measurements. The method can easily be implemented on existing computer controlled IV-measurement equipment as it is commonly used in wafer fabrication facilities. Simple measurement devices of moderate size are sufficient. The method is especially useful for semiconductor layers with low impurity concentration, as it is not influenced by channel series resistance, and correction for non-constant mobility may then be omitted.

ACKNOWLEDGEMENTS

The authors are much indebted to H. Punter and A.M.J. van Hout of Philips Elcoma, Nijmegen, The Netherlands, who fabricated the devices, and to J.C. Staalenburg, who performed the measurements.

REFERENCES

[1] Schottky, W., Z. Phys. (1942) 118

[2] Kennedy, D.P. and O'Brien, R.R., IBM J. Res. Dev. (1969) 212

[3] Verjans, J. and van Overstraeten, R.J., Solid St. Electron. 18 (1975) 911

[4] Baccarani, G. et al., Solid St. Electron. 23 (1980) 65

[5] Sze, S.M., Physics of Semiconductor Devices (Wiley, New York, 1981)

[6] Buehler, M.G., IEEE Trans. Electron Devices ED-27 (1980) 2273

[7] Masetti, G., Sevri, M. and Solmi, S., IEEE Trans. Electron Devices ED-30 (1983) 764

[8] Hendrickson, T.E., Solid St. Electron. 22 (1979) 199

[9] Ralston, A., A first course in Numerical Analysis (McGraw-Hill, New York, 1965)

[10] Wolsheimer, E.A. and Kleefstra, M., IEEE Trans. Electron Devices ED-29 (1982) 1930

SCANNING OF THE ENTIRE ENERGY GAP AT THE Si-SiO_2 INTERFACE IN MOSFETs USING THE CONDUCTANCE TECHNIQUE : COMPARISON WITH DYNAMIC TRANSCONDUCTANCE MEASUREMENTS

Hisham Haddara and Mohamed El-Sayed*

Laboratoire de Physique des Composants à Semiconducteurs,
ENSERG, 23 rue des Martyrs, 38031 Grenoble Cedex, France.

A new and accurate approach to ac conductance measurements on MOSFETs is presented. It is shown that the conductance technique can be used to study interface trap properties in the entire silicon band-gap by direct measurement on a single MOSFET. Moreover, the validity of the dynamic transconductance method is assessed by comparing its results with those obtained from conductence measurements on the same devices.

1. INTRODUCTION

In spite of the intensive efforts to study the Si-SiO_2 interface properties in MOSFETs [1,2], still there is not a reliable technique capable of providing complete information about interfce traps in the whole energy gap from direct measurement on MOS transistors. Recently, a new method has been proposed to determine interface trap properties from measurements of the dynamic transconductance of MOSFETs in weak inversion [3].

In this paper, we present an adaptation of the conductance technique [4] so that it may be applied for the characterization of interface traps directly on MOSFETs. In addition, the obtained results are compared with those provided by dynamic transconductance measurements.

2. THEORY

In this section the MOSFET ac equivalent circuit in depletion and in inversion is analyzed. Furthermore, the equivalent parallel admittance of interface traps is extracted in terms of the measured MOSFET admittance. The case of an N-channel device is considered; the P-channel case follws by analogy.

Figure (1) shows the simplified MOSFET ac equivalent circuit with the source, drain and substrate grounded. The resistance r_i is the channel sheet resistance related to the inversion charge Q_i and the effective mobility μ_{eff} by : $r_i = 1/(\mu_{eff} Q_i)$.

It is important to note that the channel resistance $R_i = r_i L/W$, limits the frequency response of the inversion layer and increases the channel time constant τ_{gc} which can be expressed as [5]:

$$\tau_{gc} = \frac{C_i (C_{ox} + C_d)}{C_i + C_d + C_{ox}} \frac{r_i L^2}{4} \qquad (1)$$

Figure (2) shows theoretical plots of τ_{gc} as a function of V_g with the channel length as a parameter calculated using field dependent μ_{eff}. It is easily seen that τ_{gc} is essentialy constant in weak inversion and decreases with increasing V_g in strong inversion. It is also, worth noting that τ_{gc} has values less than 10^{-8} s for $L < 5$ µm. This means that, for such channel lengths, the effect of the inversion layer resistance on the time response of the channel can be neglected at frequencies up to several tens of MHz.

Having developed the equivalent circuit of an MOSFET (with source, drain and substrate tied together) we are now going to discuss the exploitation of conductance measurements for the determination of interface trap properties in the entire silicon band-gap.

* Electrical Eng. Dept., Alexandria University, Alexandria, EGYPT.

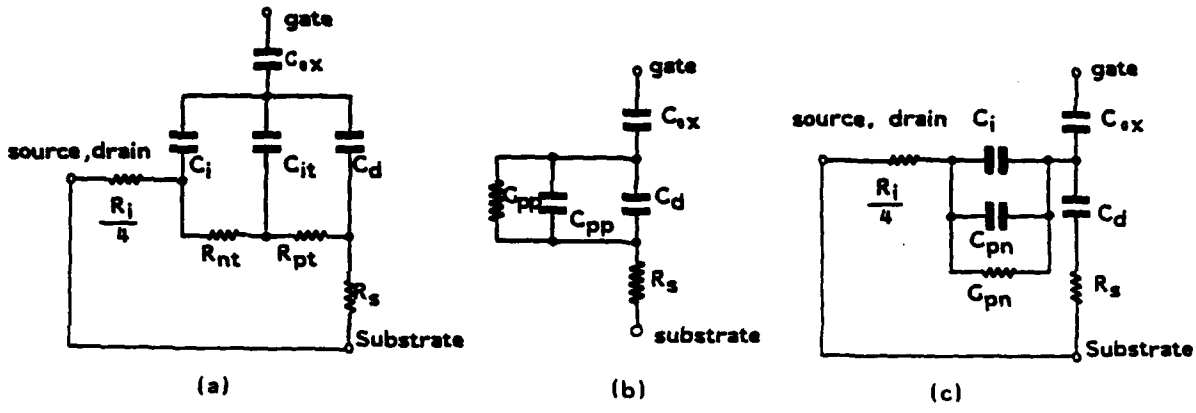


FIGURE 1

Simplified MOSFET ac equivalent circuit with grounded source, drain and substrate for : (a) Interface traps in interaction with both minority and majority carrier bands, (b) Depletion regime and (c) Inversion regime.

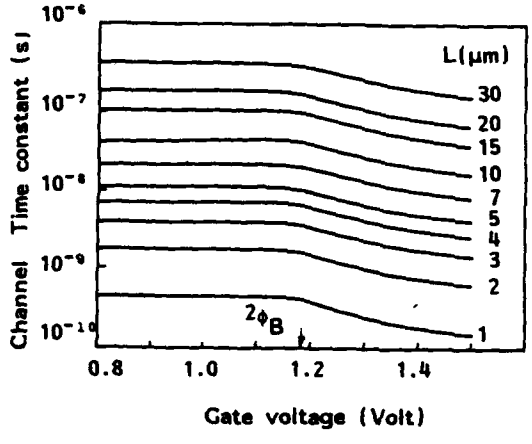


FIGURE 2

Channel time constant versus gate voltage calculated for transistors of different channel lengths and having $N_a = 10^{16}$ cm^{-3}, $t_{ox} = 25$ nm (without interface traps and fixed oxide charge).

In accumulation and depletion, the MOSFET eqivalent circuit is as shown in Fig.(1-b). G_{pp} and C_{pp} are the parallel conductance and capacitance associated with the interface traps while the resistance R_s is the substrate series resistance. We notice that the equivalent circuit in this case is exactly the same as that of an MOS capacitor and consequently, the interface traps interacting with the majority carrier band can be characterized in the same manner.

On the other hand, in the case of MOS capacitors operating in weak and strong inversion, the channel formation is very slow unless minority carrieres are supplied by an external source. Consequently, when carrying out conductance measurements, the interface traps are always obliged to interact with the majority carrier band and this requires very low measuring frequencies to study the interface traps. On the contrary, in MOSFETs, the source and drain junctions provide an enormous supply of minority carriers to form the channel. Furthermore, for negligible channel resistance (i.e. small channel lengths) the formation of the inversion layer is extremely fast. As a result of such a rapid response, the interface traps in the upper half of the silicon band-gap will easily interact with the minority carrier band in weak and strong inversion. Therefore, such traps can be characterized in a similar way as those interacting with the majority carrier band in depletion. The equivalent circuit in this case can be represented in the form shown in Fig.(1-c) with G_{pn} and C_{pn} are the interface trap parallel conductance and capacitance.

It is clear that the determination of G_{pn} and C_{pn} requires the knowledge of both R_i and C_d. The channel resistance R_i at a given value of V_g can be easily obtained from the the static $I_D(V_D)$ characteistics at very small drain voltage V_D. For determining C_d, the gate-substrate capacitance is measured with the source and drain floating. In this case the minority carriers (electrons) forming the channel are not provided by the source and drain but by the substrate itself and consequently, the channel response is very slow. In fact we have exactly the same

situation as in MOS capacitors.

Finally, the interface traps in the entire silicon band-gap can be characterized using the G_{pp}/ω and G_{pn}/ω vs. frequency plots at different gate voltages in depletion and in inversion.

3. EXPERIMENTAL RESULTS AND DISCUSSION

Conductance, capacitance and dynamic transconductance measurements have been performed on poly-silicon gate n-channel MOSFETs having substrate doping of $2X10^{16}$ cm^{-3}, gate oxide thickness of 25 nm, channel length L=3 μm and channel width W=6000 μm.

Measurements are carried out at room temperature using the HP4192A impedance analyzer at frequencies between 100 kHz and 10 MHz. At lower frequencies (50 Hz<f<100 KHz) the PAR-124A synchronous detector equiped with the PAR-184 current-voltage converter is used.

it is worth noting that the application of the conducatnce technique to VLSI MOSFETs of very small dimensions requires very sensitive measuring instruments and special precautions since the device gate capacitance is in the order of a few tens of femto Farad. It is for this reason that we have performed our measurements on transistors of large channel widths in order to increase the gate area while keeping the channel length in the range of a few microns. However, for MOSFETs of channel lengths of the order of 1 μm and gate oxide thicknesses smaller than 15 nm (typical for VLSI applications), channel widths of few hundreds of microns will be quite sufficient to allow easy capacitance and conductnce measurements.

For illustration, Fig.(3) shows the measured capacitance $C_m(V_g)$ and conductance $G_m(V_g)$ characteristics for two different frequencies obtained with the source, drain and substrate grounded. In addition to the classical conductance peak in depletion, shifting towards accumulation with increasing frequency, a second conductance peak appears in weak inversion. This second peak shifts towards strong inversion with increasing frequency and is, therefore, attributed to the interaction of the interface traps with the conduction band due to the rapid response of the inversion carriers. Furthermore, with decreasing frequency, the two peaks shift towards mid-gap voltage and are confused in a single peak.

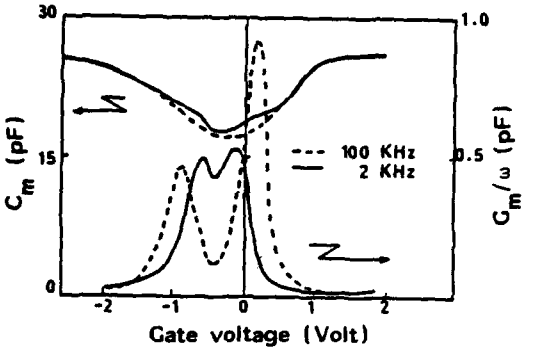


FIGURE 3

Measured capacitance C_m and conductance G_m versus gate voltage.

The interface trap density D_{it} and the time constants τ_p and τ_n can be determined from the magnitudes and positions of the maxima of G_{pp}/ω and G_{pn}/ω. In weak inversion G_{pn}/ω is obtained from conductance as well as from dynamic transconductance measurements. In the later case, G_{pn}/ω is found from the imaginary part of $1/g_m(\omega)$ and is given by [3]:

$$\frac{G_{pn}}{\omega} = \frac{I_d\, C_{ox}}{(kT/q)}\ \mathrm{Im}\,[1/g_m(\omega)] \qquad (2)$$

where I_d is the drain current at a given V_g.

Figures (4) and (5) show the interface trap density profile $D_{it}(E)$ and the time constants $\tau_{p,n}$ in the entire silicon band-gap. An excellent agreement is observed between conductance and dynamic transconductance results in the upper half of the band-gap.

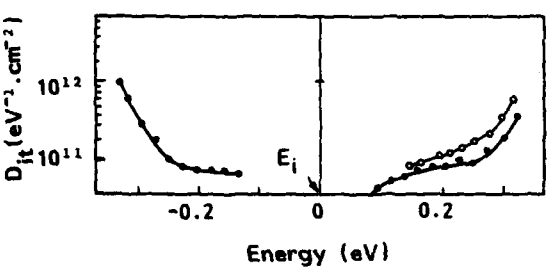


FIGURE 4

Interface trap density profiles obtained using the conductance technique (filled circles) and the dynamic transconductance method (open circles).

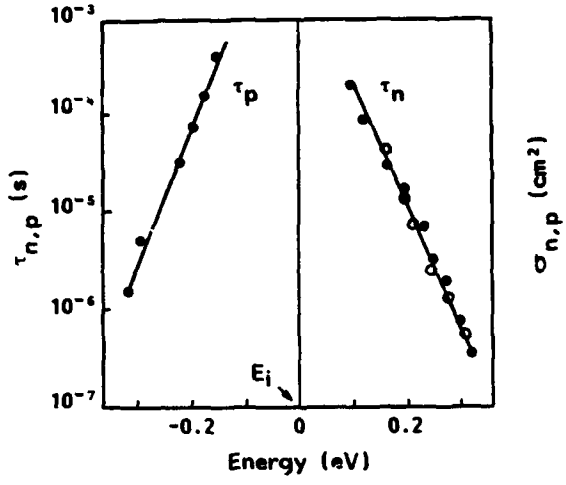


FIGURE 5

Interface trap time constants obtained using the conductance technique (filled circles) and the dynamic transconductance method (open circles).

4. CONCLUSIONS

In this paper we adapted the conductance technique, well known for MOS capacitors, to be applied to MOSFETs in inversion as well as in depletion. We have clearly illustrated that the conductance technique can be applied to study interface trap properties through the whole silicon energy gap by direct measurements on a single MOSFET. As a result, an MOSFET of a sufficiently large channel width and relatively small channel length ($L < 10$ μm) can be used as a test structure instead of on-chip monitor capacitors to study interface trap properties.

Finally, the dynamic transconductance method is proved to give the same information as the conductance technique in weak inversion. Furthermore, this method, unlike the conductance technique, can be used in the case of very small gate areas by virtue of its high sensitivity [3].

REFERENCES

[1] K.L. Wang and A.O. Evwaraye, J. Appl. Phys. 47, 4574 (1976)

[2] H. Haddara and S. Cristoloveanu, Solid-St. Electron. 29, 767 (1986).

[3] H. Haddara and G. Ghibaudo, Europhys. Conf. Abstracts, 10G, 47, ESSDERC'86, Cambridge, UK (1986).

[4] E. Nicollian and G. Goetzberger, The Bell Syst. J. 46, 1055 (1967).

[5] P. Chow and K. Wang, IEEE Trans. Electron. Dev. ED-33, 1299 (1986).

P2.2.11

HIGH RESOLUTION INTRINSIC MOS CAPACITANCE - MEASUREMENT SYSTEM

P. LECLAIRE

BULL S.A. - B2/028
Rue Jean Jaurès
78340 Les Clayes-Sous-Bois - France

Project (Spectre) ESPRIT n° 554

The difficulties of making accurate measurements of intrinsic capacitances on small geometries MOSFETs renders the modelization of these components unreliable. Recently an on-chip technique has been developped for these measurements (1). This method, however, is not versatile because it need special on-chip circuitry. This paper describes a high resolution direct on-wafer off-chip C-V capacimeter.

This apparatus allows for the validation of the SPICE intrinsic gate capacitance model by the observation of the differences in capacitances when the geometries are reduced or when the transistors are not conventionnal (ie- LDD, DDD...). It also permits the measurement of the effect of charges variation in the oxide or at the SIO2 interface on the Cgd and Cgs capacitance after hot carrier aging. In this presentation, measurements have been carried out on long and short transistors in order to test the validity of SPICE. Measurements have also been carried out before and after hot carrier stressing, and it will be shown that the capacitance characteristics are strongly affected by this stressing.

The detailed signal path is shown in figure (A). The measurement system is composed of an L.I.A and three bias sources under the control of an Apple IIE microcomputer. An I-V converter transforms the ac current signal into an amplified ac voltage. The frequency of the input signal is 10 KHZ and the amplitude is 10 mv rms. The resolution of the system is in the atto farad (1E-18 farad) range. The accuracy is less than four percent for the range 5 femtoFarad (5E-15 Farad).

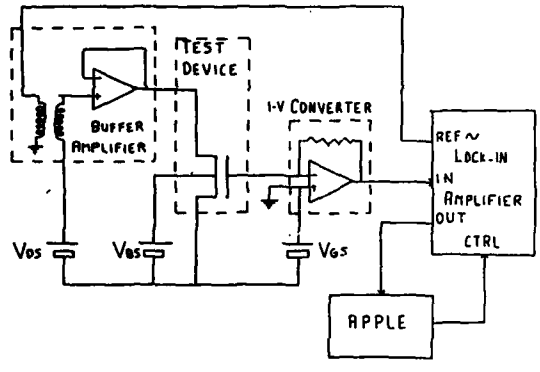


FIGURE A

The figure (B) shows the variation on the overlap capacitor versus VDS and VGS of a standard test MOSFET with channel length and width of 1.5 micron and 10 micron respectively and an oxide thickness of 280 A. The Y-axis of the graph represent a capacitor variation of 1 fF. The smallest difference between two adjacent points is 25 aF. These curves are characteristic of the moving of the depletion edge of the P-N jonction on the drain side near the SIO2 interface with the electric field.

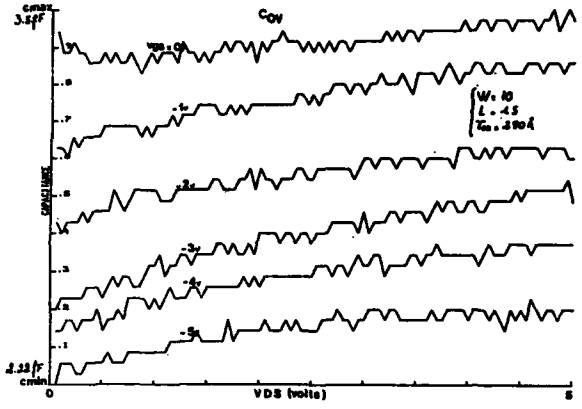


FIGURE B

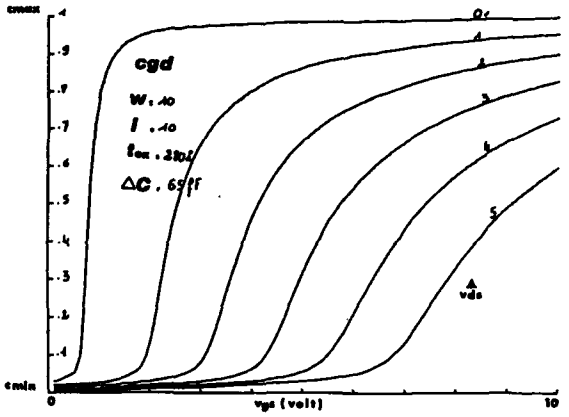


FIGURE C.2

The figure (C), (D) show the Cgs and Cgd capacitor of N channel transistors. The geometries of these device are 10/10 and 50/1.5 For comparison, SPICE simulation results are shown in figure (E). The long channel MOSFETs are well modelized but the short is not, because SPICE doesn't include two-dimensionnal effects which are characteristic of the differences observed between long and short in saturation and depletion (region I and II in the figure) ; (2), (3).

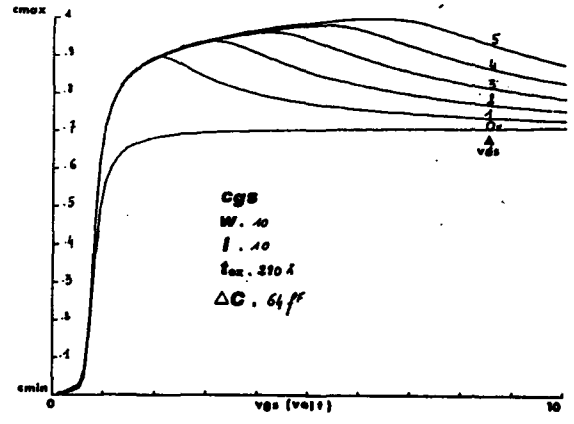


FIGURE C.1

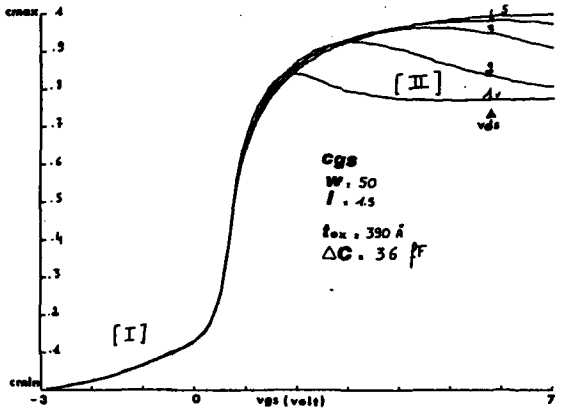


FIGURE D.1

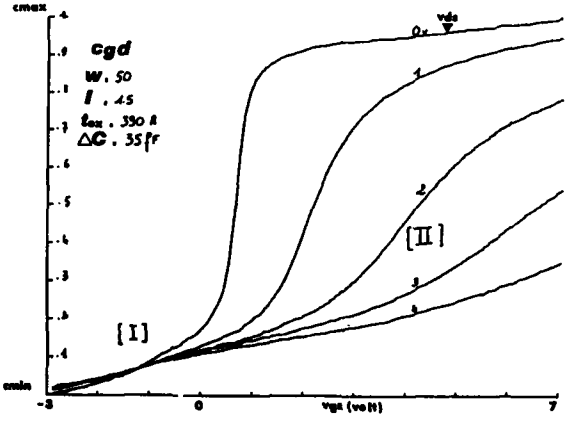


FIGURE D.2

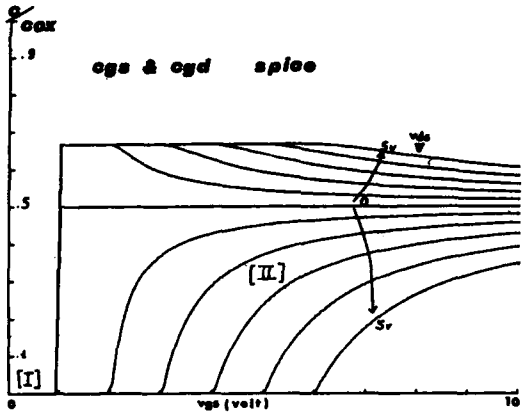


FIGURE E

The figure (F) shows the measured values of Cgd for a short and narrow channel transistor (W/L = 5/2). The oxide thickness is 390 A. We can observe, in comparison with an long and large transistor, the effect of the increasing of the electric field which is responsible of the velocity saturation effect, (4), characterized by the weakening of the gradient of the curve for high VDS and VGS. Despite the fact that the SPICE model is not representative of the small transistor (L < 2.5 micron), it nevertheless gives the worst case and is thus not critical from the circuit designer's point of view.

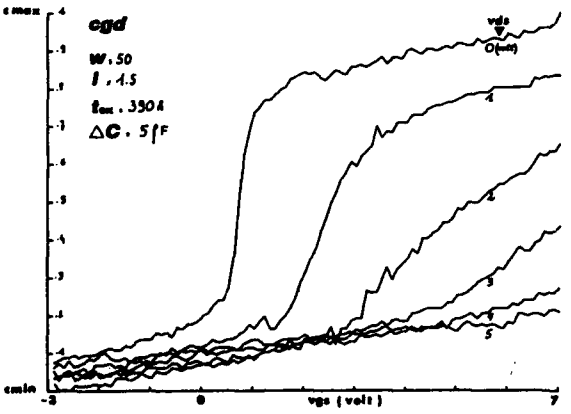


FIGURE F

We have used this system to confirm the assumption that the charge du to hot carrier is localized near the edge of the drain jonction. Figure (G) and (H) shows the curves Cgs and Cgd before and after aging. The comparison of these curves, Cgs and Cgd, before and after stressing shows outstanding events in the desertion and saturation region (I and II on he curves).

We note a difference between the curves obtained before and after aging in the desertion region. After aging, for a given Vgs (Vgs < 0) the capacitor Cgs versus VDS is increased and Cgd versus VDS is decreased due to the sheilding of the transversal field near the drain region by the localized charge which results in a better coupling between source and gate and a weakened coupling between drain and gate.

This assumption is confirmed in the saturation region both for Cgs and Cgd. For small VDS, the curves before and after aging are similar. For high drain bias, the value of the gradient of Cgs capacitance versus Vgs is increased while Cgd is lowered. In this state, the gradient is characteristic of the velocity of the carrier in the channel and therefore the value of the transversal field. The variation of the gradient for these two curves shows that the parallel surface component of the field is decreased (4) and the fact that this variation is positive for Cgs and negative for Cgd shows that the field is modified only near the drain region.

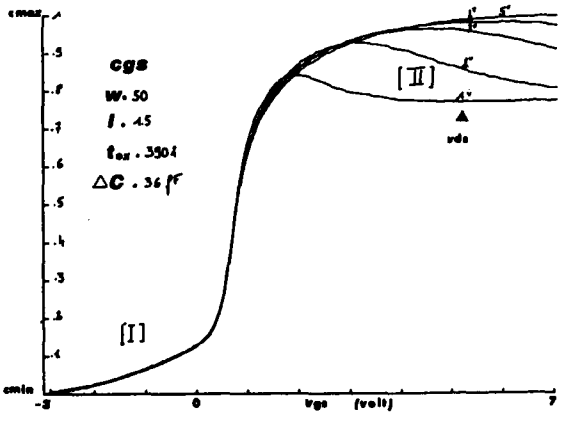


FIGURE G.1

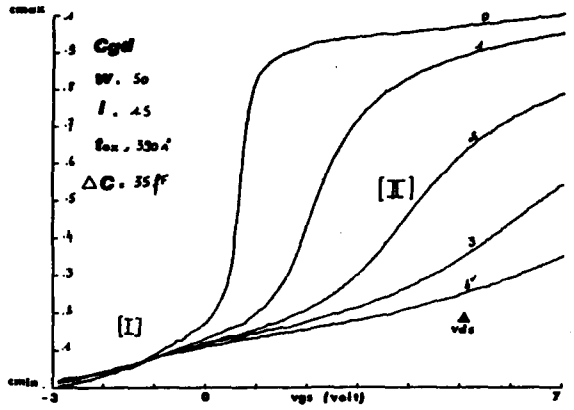


FIGURE G.2

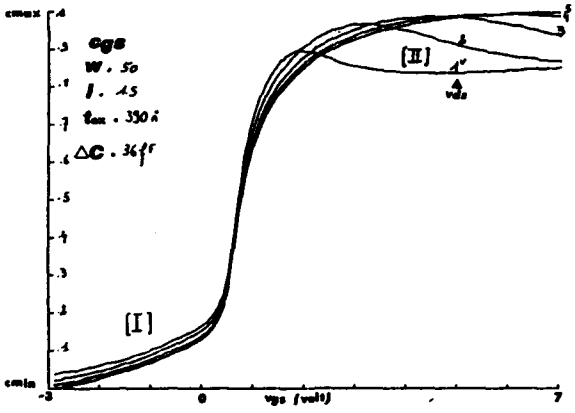


FIGURE H.1

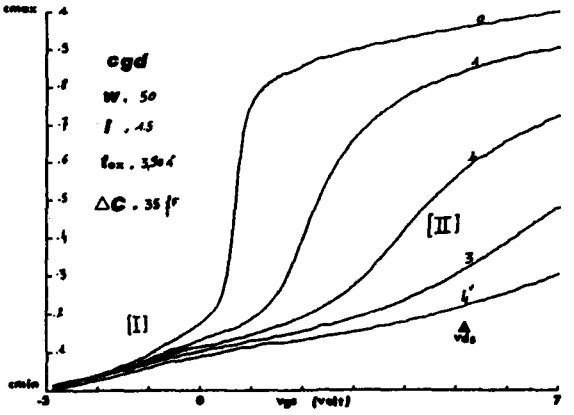


FIGURE H.2

REFERENCES

[1] A scaleable Technique for the Measurement of Intrinsic MOS Capacitance with Atto-Farad Resolution, H. Iwai & al, Solid-State circuits, vol.sc-20, feb 1985.

[2] A Measurement-based Charge Sheet Capacitance Model of Short-Channel MOSFET's for SPICE, H.J.Park & IEEE IEDM 1986.

[3] An Analytical Model fort Intrinsic Capacitances of Short-Channel MOSFET's, Bing.J.Sheu, IEEE IEDM 1984.

[4] Velocity Saturation Effect on Short-Channel MOS Transistor Capacitance, H.Iwai & al, Electron Device letters, vol.edl-6, n°3, March 1985.

EXPERIMENTAL EVIDENCE FOR DIFFERENT SATURATION VELOCITIES OF ELECTRONS IN SILICON

B.Borchert and G.Dorda

Siemens AG, Corporate Research and Development
Otto-Hahn-Ring 6
D-8000 Munich 83

This paper presents the piezoresistance effect of n-inversion layers in the hot-electron regime. The measurements were performed on short-channel n-MOSFETS both at 77 and 300K. From the experimental data clear evidence is obtained for different saturation velocities of electrons in Si depending on the occupation of the subbands. Including this effect good agreement between theory and experiment mainly at 300K is achieved.

1. INTRODUCTION

Nowadays hot-electron effects play an important role in VLSI-devices. Because of their negative influence on transistor operation they are the object of extended research activities (1). In the present paper we use piezoresistance effect (PE) measurements on short-channel MOSFETS to study their features. The PE is caused by a repopulation of electrons among the subband levels of the different valleys under a uniaxial stress (2). Since the PE enables us to get information about the electrons in the different valleys and to study quantum phenomena of the inversion layer it is a promising tool especially in the hot-electron regime.

2. THEORY

In the theory of the PE the hot-electron effects like velocity saturation and electron temperature are included. Here also the inhomogeneous electric field distribution within the transistor is incorporated. For the subband quantization of the n-inversion layer we used the triangular potential wall approximation, where the subband levels as a function of the electric surface field can be expressed analytically (3). The appropiate equations for the PE are listened in (4,5).

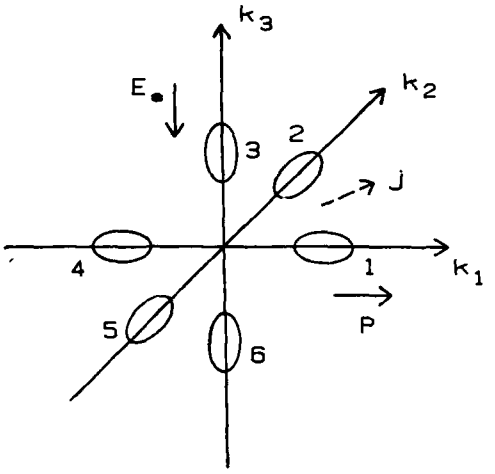


FIGURE 1: Diagram of constant energy surfaces in k-space for n-Si, the experimentally used directions of the surface field E_s in (001), of the uniaxial stress P in (100) and current j in (110) are also given

If we apply an uniaxial mechanical stress which is accompanied by an energy shift among the six conduction band valleys we get a repopulation of the electrons between the different valleys. Because of the different conductivity masses for the subbands, the electrons undergo

a mobility change during the repopulation and consequently the total mobility of the inversion layer changes. For a tension stress p in the experimental PE-configuration, see Fig.1, the valleys 1 and 4 will be lifted, while the other four valleys 2,3,5 and 6 will be lowered. Considering the conductivity masses of the valleys for an electric current j in (110)-direction the total mobility will be enhanced. The magnitude of the PE in quantizised inversion layers depends on the subband spacings and on the mobility differences between the subbands, as well as on the electron temperature. Therefore we can expect a gate and drain voltage dependence of the PE in short-channel MOSFETS.

3. EXPERIMENTAL

The measurements were performed on short-channel MOSFETS with channel lengths ranging from 1 to 30μ. By sawing probe stripes from the wafer, fixing one end and bending the other end up or down we were able to apply an uniaxial mechanical stress to the transistor. With the knowledge of the amount of the bending and the geometrical dimensions the corresponding uniaxial stress can be calculated (6). For the probe orientation see Fig.1.

4. RESULTS AND DISCUSSION

4.1. 300K-Results

Fig.2a,b shows the gate and drain voltage dependence of the PE at 300K. From the low-field (Vd =.1V) to the high-field regime (Vd >2V) we see a remarkable decrease of the experimental PE curves. This can be attributed pre-dominantly to the decreasing mobility differences of the electrons in the different valleys with increasing lateral electric field. The theoretical PE curves are in very satisfactory agreement with the experiment, see also Fig.2a,b. For the ohmic mobility we used a three subband model of Moore and Ferry (7). The

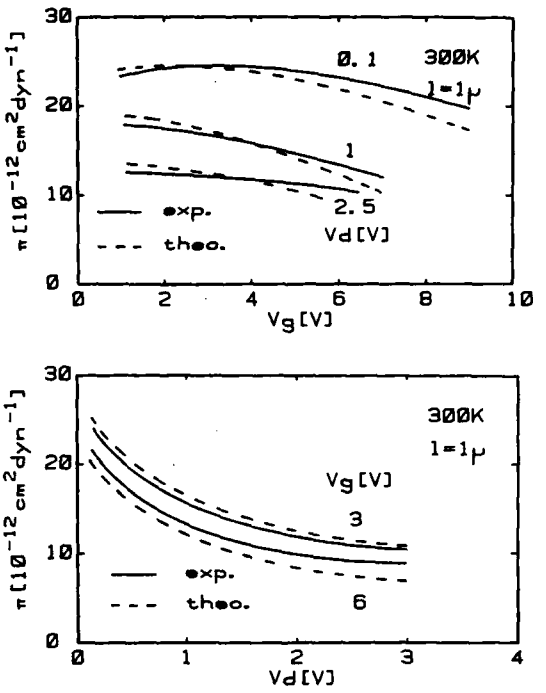


FIGURE 2: Experimental and theoretical gate voltage Vg (a) and drain voltage Vd (b) dependence of the PE coefficient π for a 1μ transistor at T=300K

fitted values for the hot-electron parameters are in good agreement with the literature (4,5).

4.2. 77K-Results

Fig.3a,b demonstrate the gate and drain voltage dependence of the PE at liquid nitrogen temperature. Here we see a much faster decrease of the PE with increasing drain voltage than at room temperature. This is a consequence of the fact that at 77K the ohmic mobility is several times larger than at 300K and therefore the velocity saturation sets in at smaller electric fields. By looking to the drain voltage dependence, we see surprisingly a saturation behaviour of the PE at high drain voltages. Since in the corresponding electric field regime the electrons have reached the saturation velocity, the theory would predict here a vanishing PE, if the saturation velo-

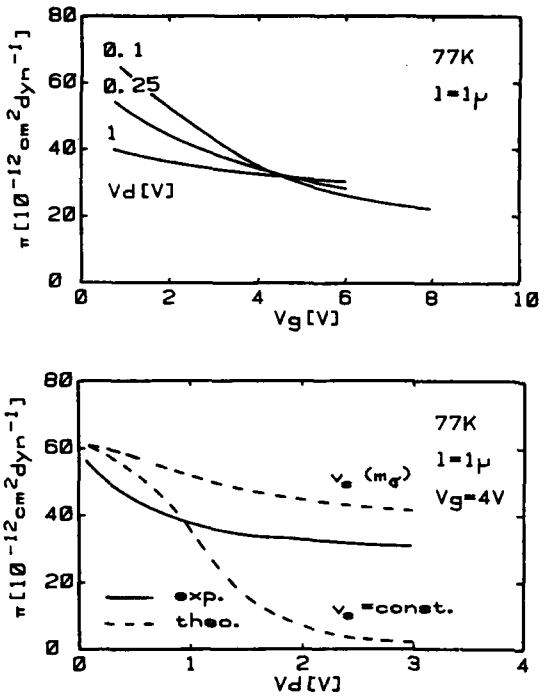


FIGURE 3: Experimental and theoretical Vg (a) and Vd (b) dependence of at T=77K

city v_s is a constant for all valleys, see Fig.3b. Because this is in clear contradiction to the experiment, we can deduce different saturation velocities for the electrons in the different valleys.

A model of Shockley (8) gives a possible explanation, which describes a concuctivity mass dependence of v_s. Within this model the electrons can reach saturation velocities from $5*10^6$ to $1.2*10^7$ $cmsec^{-1}$ varying with subband population. Based on this model the corresponding theoretical PE curves give an outstanding agreement with the experiment both at 77 and 300K, see Fig.2a,b and 3b.

5. CONCLUSIONS

It was demonstrated that the PE is a powerful tool for studying hot electrons in short-channel MOSFETS. With the help of the PE we have shown convincingly for tho first time that the value of the saturation velocity for electrons in silicon is not a constant as usually assumed. It depends on the population of the subbands, i.e. in particular on the doping of the Si substrate, the gate voltage and temperature of the electrons in the hot-electron regime. This circumstance has to be considered in the simulation program of small size MOSFETS.

ACKNOWLEDGMENT

The authors would like to thank M.Henzler and U.Bürker for her continous encouragement.

REFERENCES

(1) P.E.Cotrell, R.R.Troutman and T.H.Ning, IEEE Trans. Electron Devices,ED-26,520,1979

(2) G.Dorda, J. Appl. Phys., 42, 2053, 1971

(3) F.Stern, Phys. Rev., 5, 4891, 1972

(4) B.Borchert, M.S. thesis, University of Hannover Germany, 1987

(5) B.Borchert and G.Dorda, to be published in IEEE, Trans. ED

(6) G.Joos, Lehrbuch der theo. Physik (Akadem. Verlagsgesellschaft, Leipzig 1945), p.155

(7) B.T.Moore and D.K.Ferry, J. Vac. Sci. Technol., 17, 1037, 1980

(8) W.Shockley, Bell System Tech. Journal, p.990, Oct. 1951

P2.2.13

PERMEABLE BASE TRANSISTORS WITH HIGH TRANSCONDUCTANCE BUILT ON LPVPE-GROWN EPILAYERS

A. Gruhle, L. Vescan, * H. Beneking

Institute of Semiconductor Electronics, Sommerfeldstr., D-5100 Aachen, W. Germany
*Present Address: University of Michigan, EECS Dept., Center of High Speed Microelectronics, Ann Arbor, MI 48109, USA

There has been considerable interest in permeable base transistors (PBTs) for use in high-speed circuits. A gain of 11dB at 40 GHz has been reported /1/. An important future application could be the three-dimensional integration by stacking several devices on top of each other. PBTs consist of a base-grating embedded in semiconducting material. Using MBE the base metal or silicide may be burried under an epitaxial layer. In this paper however we present etched-groove silicon PBTs that have been patterned with reactive ion etching. This approach reduces input capacitances but has the disadvantage of non-planar surfaces.

Starting material are (100)-1 mOhmcm n^+ substrates with a 2μm thick, $4x10^{16}$ cm^{-3} n-epilayer grown by LPVPE at 820°C. The top 100 nm are n^+ doped for good ohmic contacts to the TiPtAuPt source-metallisation which is patterend by lift-off. The metal acts as a mask in the subsequent RIE where 0.7μm silicon are removed in an SF_6 plasma resulting in a slight undercut. This is neccessary to form a discontinuous metal film during the following 100 nm Pt gate metallisation. Finally, the surface is planarized with PIX-1400 polyimide, a part of which is removed in a carefully controlled subsequent O_2 plasma etch. The top side contacts are thus revealed (as can be seen in Fig.1) and reinforced with an additional Au-pad. A AuSb metallisation on the back side of the wafers forms the drain contact.



Fig.1: Polyimide-covered PBT with revealed top contacts

A series of PBTs has been fabricated with varying channel width (0.5μm - 2.0μm), groove depth (0.2μm - 1.5μm) and device length (50μm - 4000μm). Dual-gate transistors for mixer applications have also been built /2/. In the micrograph of Fig.2 two PBTs (1μm and 1.5μm channel size) can be seen.

Threshold voltages vary between -0.5V and -7V. The gate Schottky diodes have ideality factors less than 1.1 and breakdown occurs at about -10V. Fig.3 shows the typical IV-characteristics of a 0.5μm-PBT. The maximum obtained transconductance is 62 mS/mm, the highest value ever reported for Si PBTs. This is achieved by the optimized $n^+n^-n^+$ epilayer structure that permits very low source, drain and contact resistances. For the same reason there is almost no change in the IV-curves when the top contact is used as drain.

PBTs suffer from a large output conductance as a result of drain voltage influence on drain current. Presently two-dimensional numerical simulations are beeing performed to evaluate the influence of different gate configurations on the output conductance. RF measurements are presently performed on large-area PBTs (channel length 1000μm to 4000μm) with low impedance adjusted to 50 Ohm microstriplines.

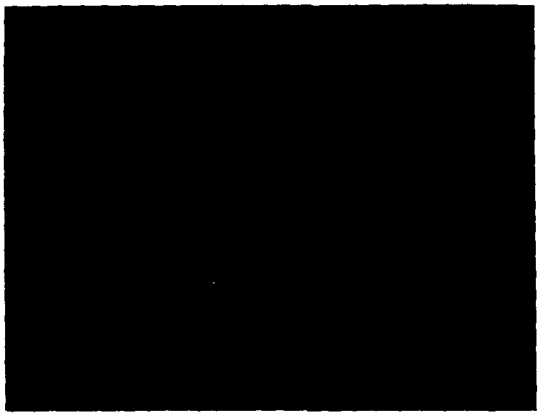


Fig.2: Micrograph of two PBTs. Gate pad size is 60μm x 30μm.



Fig.3: Typical IV-characteristic of a 60μmlong PBT. Gate offset is +0.5V.

REFERENCES

/1/ R.Actis et al., IEEE Vol.EDL-8, No.2, 66, 1987

/2/ A.Gruhle, L.Vescan, H.Beneking, Electr. Lett., Vol.23, No.9, 447, 1987

PREDICTION OF SOFT ERROR RATE OF 4 Mbit DRAM

W.H. Krautschneider and W. Meyberg
Siemens AG, Central Research and Development
Otto-Hahn-Ring 6, D-8000 Munich 83, FRG

A method has been developed for estimation of soft error rate of memory chips. At special test structures which are as simply designed as possibly the charge collection induced by alpha strikes is measured. From these data the soft error rate can be calculated.

1. INTRODUCTION

Soft errors caused by alpha particle impacts are a major concern for the reliability of dynamic memories. For that reason it is important to have means for predicting the soft error rate as a function of process parameters in an early stage of the design of dynamic memories.
The method proposed is based on charge collection measurements taken at test structures equivalent to cell and bitline nodes but without access transistor, as memory operation is not required for charge collection experiments. Since these test structures require less process steps the alpha particle sensitivity can be determined some time before fully functional memory chips are available.

Dynamic memories of the 4 M generation use trench cells for achieving the necessary cell capacitance on a scaled down area. Adjacent trench cells can be shortcircuited by alpha tracks [1]. This effect has been experimentally verified and also taken into consideration for soft error rate prediction.

2. MEASUREMENT OF COLLECTED CHARGE INDUCED BY ALPHA PARTICLE STRIKES

An alpha particle striking silicon generates electron-hole pairs along its track which are partly collected from the storage cell by carrier drift [2,3] and diffusion [4,5].
The basic measurement setup to determine the charge which is transferred when an alpha particle hits the device shows Fig. 1. An active probe tracks the voltage swing $U_m(t)$ at the node under test which was biased to 5 V by a voltage divider which is dimensioned that its time constant is much longer than the duration of the alpha signal. The signal is preamplified and digitized by a fast analog-digital converter.

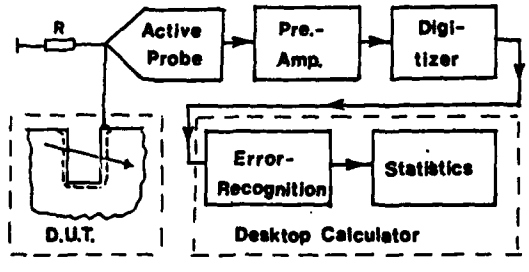


Fig. 1 Basic measurement setup

The digitized signal is stored and processed using a desktop calculator which first checks the input signal for error recognition and suppression. Then the digitized input signal $U_m(t)$ is integrated to determine the transferred charge

$$Q_m = \int_0^{t_o} U_m(t)/R \, dt$$

The wiring of the test structures is arranged in such an order that only cells with a distance of about 30 µm from each other are contacted. Thus the charge spreading over to neighbouring test structures can be neglected and the charge transferred to just one node is measured.
When due to the necessary contact the area of the cell test structure is a bit larger than the actual cell node some corrections have to be made. Numerical calculations show that the charge collected by diffusion is proportional to the square root of the collection area at small structures [5]. Interpolation between the charge collected by cell and bitline test structures yields for the corrected cell node charge

$$Q_{cor} = (\sqrt{A_c} - \sqrt{A_{bl}})/(\sqrt{A_{ct}} - \sqrt{A_{bl}})*$$
$$*(Q_{ct} - Q_{bl}) + Q_{bl}$$

The subscripts of the area A and charge Q - c, bl, and ct - denote the cell, bitline, and cell test structure, respectively.

Trench cells collect additional charge when alpha particles traverse two adjacent cells with one of the cells in state "1" and the other in state "0" [1]. To measure the charge transfer between trench cells test structures were used consisting of two grids of diffusion areas with trenches. Double strikes cause a current flow between the affected trench cells that leads to a positive voltage drop across the resistor R_2 (Fig. 2) from which the additionally collected charge can be determined. This charge has to be multiplied by the probability that a double strike occurs, which can be obtained from geometrical calculations, and added to the charge collected by the cell test structure.
A 100 nC radium source has been used for irradiation.

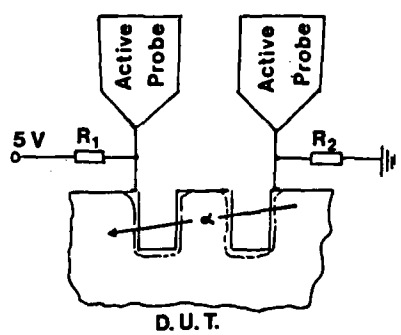


Fig. 2 Measurement setup for determination of additionally collected charge when an alpha particle traverses two adjacent trench cells

3. DETERMINATION OF CHARGE TRANSFER FUNCTION

From numerous measurements the distribution of the transferred charges N(Q) is obtained. Subsequent summation and normalization gives the distribution function F(Q) of charge collection

$$F_k(Q_k) = \sum_{i=1}^{k} N_i(Q_i) \Big/ \sum_{i=1}^{n} N_i(Q_i) \qquad k \le n$$

with the total number of alpha events n.
Completion and rearranging leads to the equation

$$\sum_{i=k+1}^{n} N_i(Q_i) \;/\; \sum_{i=1}^{n} N_i(Q_i) = 1 - F_k(Q_k) =$$

$$= P_{k+1}(Q_{k+1}) \equiv P(Q)$$

The function P(Q) describes the probability for the occurence of an alpha strike that tranfers a charge Q to the test structure which is larger than or equal to Q_{k+1}. Setting Q equal to the critical charge Q_c which is required for a reliable operation of the memory cell the function P(Q) gives the probability that an alpha strike causes a soft error (Fig. 3).

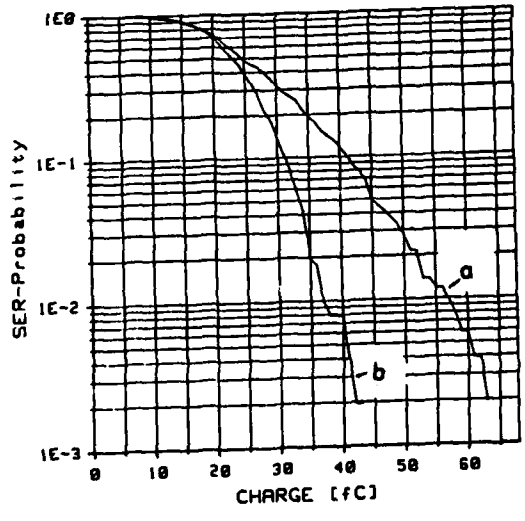


Fig. 3 Soft-error probability function for alpha particles exceeding a charge threshold
a) trench cell, b) bitline

4. CALCULATION OF SOFT ERROR RATE

Following May and Woods [6] the soft error rate (SER) of memory chips is proportional to the product of the area which is sensitive to alpha particles and the alpha flux Φ_α

$$(1) \quad SER \sim A_c * \Phi_\alpha .$$

A_c denotes the horizontal projection of the pn-junction space charge area. Proportionality factor is the sensitivity factor S of the circuit under test. It is defined

$$(2) \quad S = \sum P(E_i) * N(E_i)$$

The function $P(E_i)$ describes the probability that an alpha particle of the energy E_i causes a soft error and $N(E_i)$ is the energetic frequency distribution. Since the physical reason for the occurence of a soft error is the collection of charge which exceeds the critical charge, the product $P(E_i)*N(E_i)$ can be replaced by the charge transfer function P(Q) which can be obtained with much less effort than $P(E_i)$ and $N(E_i)$.
Inevitable fluctuations in transistor geometry and technology can lead to different threshold voltages of the sense amplifier transistors. This threshold voltage difference ΔVt reduces the critical charge of the storage cell for

$$(3) \quad \Delta Q_c = (C_c + C_{bi}) * \Delta Vt$$

which causes an increase in the soft error rate. At pairs of MOS transistors which were equivalent to the sense amplifier transistors the probability density of the threshold voltage difference g ΔVt has been determined.
Using eqn. (1) to (3) yields

$$SER(Q_c) = \Phi_\alpha * N(A_c \sum_{i=1}^{i_{max}} g\Delta Vt_i) * P_c(Q_c - \Delta Q_{c,i})$$

$$+ 0.5 * A_{bl} * s \sum_{i=1}^{i_{max}} g\Delta Vt_i * P_{bl}(Q_c - \Delta Q_{c,i})$$

with the sensitive area of the cell node A_c and bitline node A_{bl} and the number of nodes N. The factor s stands for the quotient of the time during which the node is floating to cycle time multiplied by the fraction of active nodes to the total number of nodes. For the bitlines it has to be taken into account using a factor 0.5 that either the bitline or the reference bitline is sensitive to alpha strikes.

Fig. 4 shows the soft error rate in dependence on the critical charge.

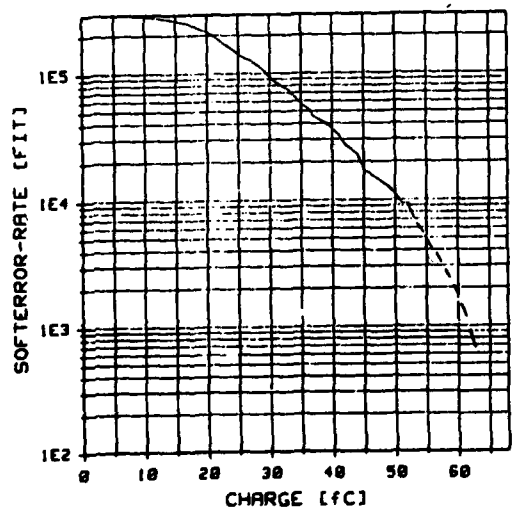


Fig. 4 SER in FIT (soft errors in 10^9 hours) as a function of critical charge at an alpha flux of 1000 particles/($cm^2 * 10^6 h$)

5. CONCLUSION

The charge collected by trench cells induced by alpha particle strikes including the charge transfer, when an alpha particle shortcircuits two adjacent trenches, has been measured and evaluated. The data are helpful for a cell design to meet the required standard of soft error rate.

ACKNOWLEDGMENT

This report is based on a project which has been supported by the Ministry of Research and Technology of the FRG under the contract no. NT 2696. For the contents the authors are responsible alone.

The provision of computer software by H. Benzinger is gratefully acknowledged.

REFERENCES

[1] J.-S. Chern, P. Yang, P. Pattnaik, and J.A. Seitchik, "Alpha-particle induced charge transfer between closely spaced memory cells", IEEE Trans. Electr. Dev., ED-33, p. 822, 1986.

[2] C. Hu, "Alpha-particle-induced field and enhanced collection of carriers", Electr. Dev. Lett., EDL-3, p.31, 1982.

[3] E. Takeda, K. Takeuchi, E. Yamasaki, T. Toyabe, K. Oshima, and K. Itoh, "The scaling law of alpha-particle induced soft errors for VLSI's", IEDM Techn. Digest, p. 542, 1986.

[4] S. Kirkpatrick, "Modeling diffusion and collection of charge from ionizing radiation in silicon devices", IEEE Trans. Electr. Dev., ED-26, p. 1742, 1979.

[5] K.W. Terrill, C. Hu, and A.R. Neureuther, "Computer analysis of the significance of surface boundary conditions on the collection of alpha-induced charge", Sol.-State Electr., vol. 26, p. 15, 1983.

[6] T.C. May and M.H. Woods, "Alpha-particle-induced soft errors in dynamic memories", IEEE Trans. Electr. Dev., ED-26, p. 2, 1979.

THE EFFECT OF THE PROXIMITY OF THE BIRD'S BEAK ON AGING OF THE THIN OXIDE BY HIGH FIELD CURRENT STRESS

by J.-C. MARCHETAUX, B. DOYLE and A. BOUDOU

BULL S.A., Ave Jean-Jaures, 78340 Les Clayes sous Bois, FRANCE.

1. INTRODUCTION

The quest for a greater integration in VLSI circuits has lead to a diminution not only of the length of MOS channel with its inherent problems (hot electrons, etc...), but also of the width. It is thus very important to know the quality of the transition (bird's beak) zone between thick and thin oxides and its influence on the quality of the neighbouring thin oxide.

We have shown [1] that during the VLSI fabrication process, the growth of the field oxide induces a large quantity of active interface states close to the active thin oxide zone under the bird's beak region (BB region).

In this presentation we show that as well as above, the proximity of the BB zone induces a dramatic accelerating aging on the active thin oxide zone when subjected to high field current stress.

2. EXPERIMENTAL

2.1. Samples

To do this we compare the aging of two capacitor structures: a thin oxide (standard) capacitor and a capacitor with a maximized perimeter/surface ratio (BB capacitor). For the latter the maximization is obtained by the parallel connection of 6400 cells of 10*10 μm^2, each of these cells consisted of a square of thin oxide of dimensions 5*5 μm^2 surrounded by the field oxide.

The samples used came from a 2 micron NMOS process with a dry gate oxide of 270 A and 400 A thichnesses, a P-type substrate and without threshold implant. The BB zone derives from a standard LOCOS process.

2.2. Aging and characterization systems

The samples are aged by electron injection in the Fowler-Nordheim regime (FN injection), from the gate to the SiO2/Si substrate interface (negative gate voltages) and in the constant current mode. In the BB structure, in spite of the changing oxide thicknesses, the injection is localized to the gate oxide i.e. to the point of thinnest oxide. Thus, any observed difference between this test structure and the standard structure stressed in the same injection density conditions will have been due to the bird's beak influence on the thin oxide zone and not to damages created in the thicker oxide of the BB region.

The characterization is carried out by the quasistatic C-V method. The modification of the method to be able to be applied to BB capacitors which have various oxide thicknesses and dopings is described in detail elsewhere [1]. To make the comparison between the two types of capacitor, at each characterization of BB samples, the field and BB capacitive components are removed by computation. For this the assumption is made that these two components are not modified by the various injections. This quite reasonable because the FN injection is localized to the gate oxide.

3. RESULTS

3.1. Aging

The curves of the evolution of the injection voltage as a function of time are given in Figures 2 (270 A) and 3 (400 A). While the

standard capacitor curves are nearly horizontal we observe for the BB structure an important modification of the curves: at the beginning there is a clear compensation of the positive charges by electron trapping ($\Delta V_{inj}<0$) which is then rapidly compensated then exceeded by a large generation of additional positive charges ($\Delta V_{inj}>0$). Moreover there is a much larger divergence between the different aging curves for the BB than for the standard samples. These two differences show that the proximity of the BB region induces a larger hole trap density in the region near the gate injecting electrode. This suggests that the centroid of trapped charge is closer to the gate in the 270 A case than in the 400 A case.

For the quasistatic C-V measurements, using the intermediate step described in 2.2. for BB samples, we obtain the profile of the interface state density (N_{ss}). The results are given in figure 4 (standard) and 5 (BB) for 270 A. Similar results are obtained for 400 A. We observe the presence of a peak generally related to dangling bonds of trivalent Si atoms [2]. In our case, the position in the forbidden gap of these states is about $E-E_v=0.6$ eV for the two structures.

On comparing the magnitude of N_{ss}, we see that , in the case of the BB samples, this magnitude is about 5 to 10 times larger than the standard structure one. This holds for the peak at 0.6 eV as well as for the background trap density.

Thus, it is clear that the proximity of the BB region, although not modifying qualitatively the type of interface traps generated under FN injection, induces in the active zone of the device a dramatic acceleration of the interface trap generation.

Figures 5 and 6 represent the variation of the flat-band voltage as a function of the injection dose, determined from the quasistatic measures. While the standard sample curve has the saturated behaviour at high Q_{inj} well known for this type of aging and for these thicknesses (>250 A)[3], for BB structure a non saturated behaviour is clearly seen: nearly linear for 400 A and a saturation followed by a increase for 270 A. This allows us to suppose that there is a continuous generation of hole traps during the FN injection as is seen for the very thin oxide thicknesses (<100 A) [4].

3.2. Relaxation

We have compared the relaxation of the two kinds of structures after they have been submitted to a high FN injection (N_{inj}=1E17 cm-2 and J_{inj}=-5E-5 A/cm^2).

Figures 7 and 8 show this time evolution of N_{ss}. We see immediately that, while there is a slightly annealing of the N_{ss} profile for the standard sample, there is a progressive increase of N_{ss} for the peak at 0.6 eV as well as for the defect background. The evolution in time follows an $A*t^n$ law, with n=0.23. The N_{ss} can be seen to have increased by a factor of three during this time.

4. CONCLUSION

We have shown that during a FN injection, the proximity of the bird's beak region induces a dramatic increase of the interface state density in the active gate zone, a generation of positive charge closer to SiO_2/gate interface, and a non-saturation behaviour of V_{fb} variation. Moreover, in place of an annealing, the relaxation is characterized by a progressive increase of N_{ss} which follows time power law in the limit of time taken.

This work is supported in part by the E.E.C. ESPRIT (SPECTRE) project, N° 554.

REFERENCES

[1] J.-C. Marchetaux, to be published, Solid State Electron.

[2] N.M. Johnson et al., Appl. Phys. Lett. 43 (6), p. 563 (1983).

[3] D.J. DiMaria, Proc. Int. Topical Conf. on Physics of SiO2 and its Interfaces, Yorktown Heights, New York, 1978.

[4] M.S. Liang and C-M. Hu, IEDM Proceedings, p. 396 (1981).

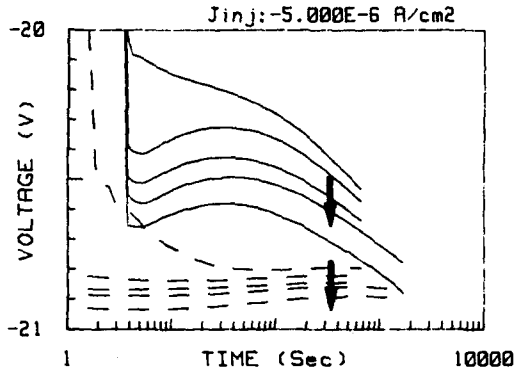


FIG. 1 Injection voltage curves at constant current density J_{inj}. Full lines: BB samples, dashed lines: standard samples. N_{inj} is varied from 1E16 to 1.6E17 cm-2 in the direction of the arrows (cf. Fig. 5). The thickness is 270 A.

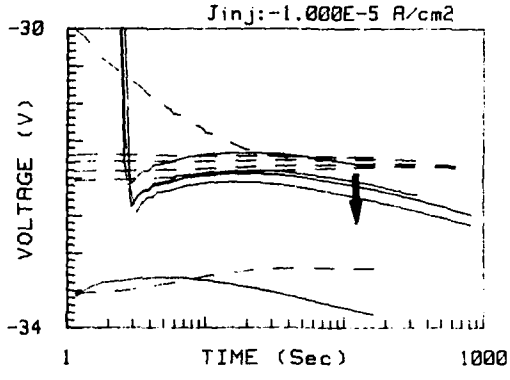


FIG. 2 Injection voltages curves at constant current density J_{inj}. Full lines: BB samples, dashed lines: standard samples. N_{inj} is varied from 1E16 to 1.9E17 cm-2 in the direction of the arrow (cf. Fig.6). The thickness is 400 A. For the two curves at the bottom: J_{inj}=5E-5 A/cm^{-2}.

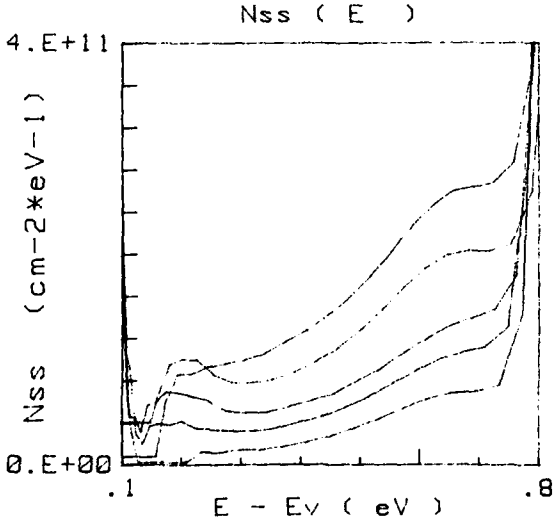


FIG. 3 Interface state density profile obtained from quasistatic C-V measurements after each aging (cf. Fig.1 and 5 for injection conditions) for the standard sample of 270 A thickness.

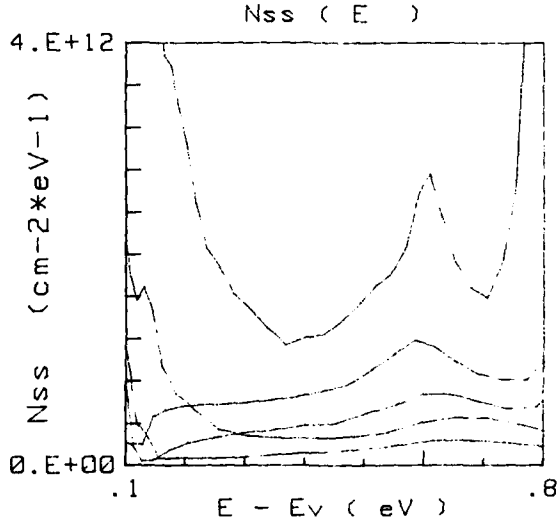


FIG. 4 Interface state density profile obtained from quasistatic C-V measurements after each aging (cf. Fig.2 and 6 for injection conditions) for the BB sample of 270 A thickness.

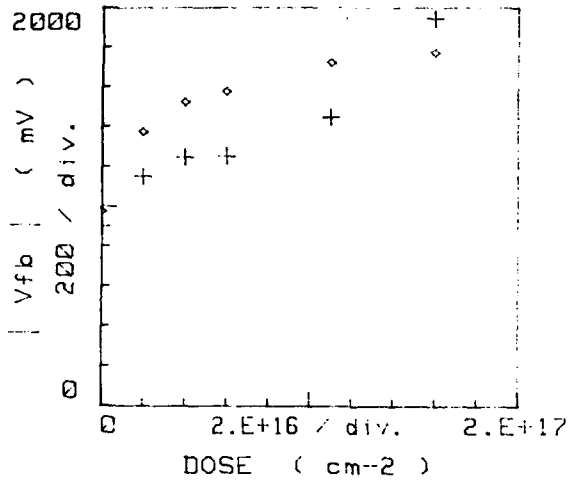


FIG. 5 Flat-band voltage curves as a function of injection dose. Diamonds: standard sample; crosses: BB sample. the thickness is 270 A.

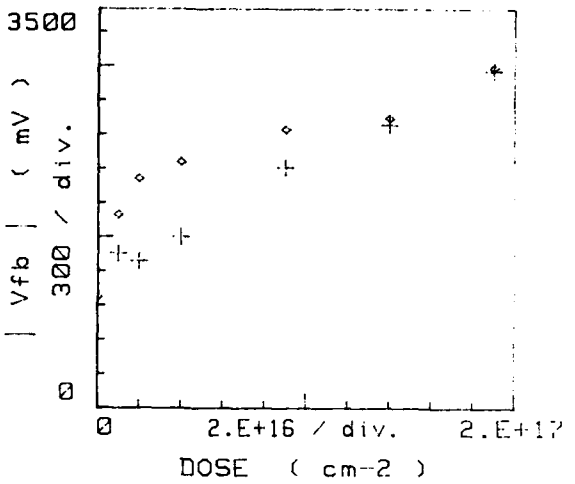


FIG. 6 Flat-band voltage curves as a function of injection dose. Diamonds: sandard sample; crosses: BB sample. The thickness is 400 A.

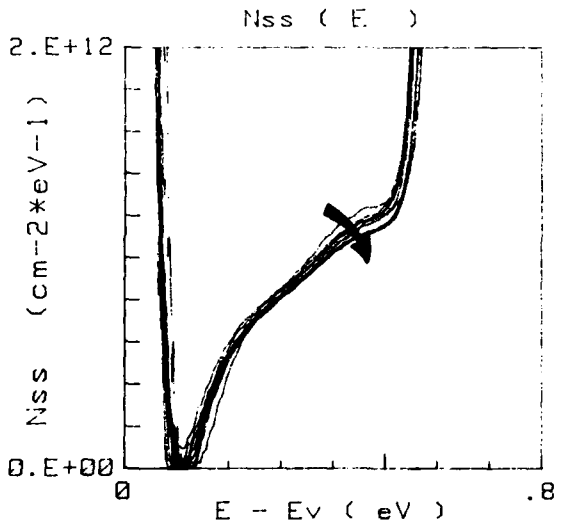


FIG. 7 Evolution of the interface state density during the relaxation of the standard sample with a 400 A thickness. Increasing is indicated by the arrow.

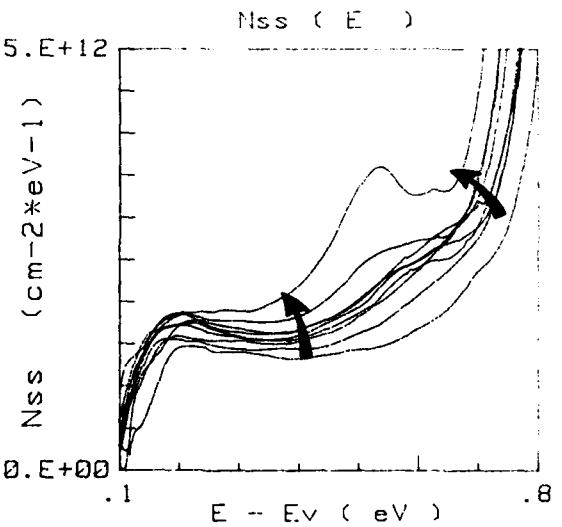


FIG. 8 Evolution of the interface state density during the relaxation of the BB sample with a 400 A thickness. Increasing time is indicated by the arrows.

SIMULATION OF STRESSED N- AND P-CHANNEL MOSFET'S: FIXED OXIDE CHARGES AND FAST INTERFACE STATES

A. Schwerin, W. Hänsch and W. Weber

SIEMENS Corporate Research and Development, Microelectronics
Otto-Hahn-Ring 6, 8000 München 83, FRG

MOSFET's stressed at high drain voltages show shifts in the device characteristics. We study the effect of the stress-induced damaged region on the device characteristics for conventional n- and p-channel devices. We give a consistent description of degradation effects on the drain current in the subthreshold and the pentode region of the MOSFET as well as on the substrate current. In that way it is possible to decide between several models which might be appealing in one or the other regime.

INTRODUCTION

Shrinking device dimensions in modern VLSI circuits leads to increasing electric fields. This causes considerable shifts in the device characteristics. General agreement exists that this degradation is due to some localized oxide charges caused by hot carrier injection in the high field region. Thus we compare experimental data with results obtained with a 2-d device simulator which allows for charges at the Si-SiO_2 interface.

DEGRADATION OF N-CHANNEL MOSFET's

a) Experimental results:

2µm-NMOS n-channel devices with a gate oxide thickness of 42nm and conventional arsenic source-drain implantations were used. In figures 1-4 we show the different features for a typical device degradation after $5x10^4$s of stress with V_D=8V and V_G=3V.

b) 2-D simulation:

For our calculations we use a modified version of the 2-D device simulator MINIMOS 3 /1/.

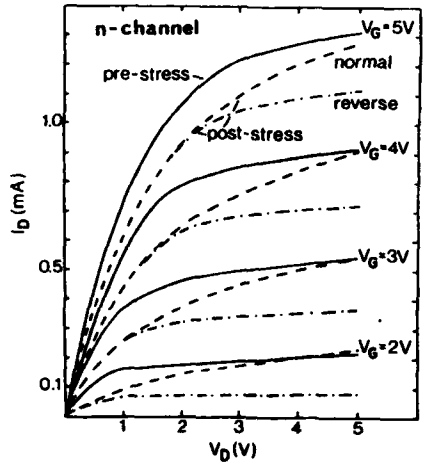


FIGURE 1
Drain current vs. drain voltage - experiment; solid lines (pre-stress), dashed lines (post-stress normal); dashed-dotted lines (post-stress reverse); stress conditions: V_G=3V, V_D=8V, t_{stress}=50000s

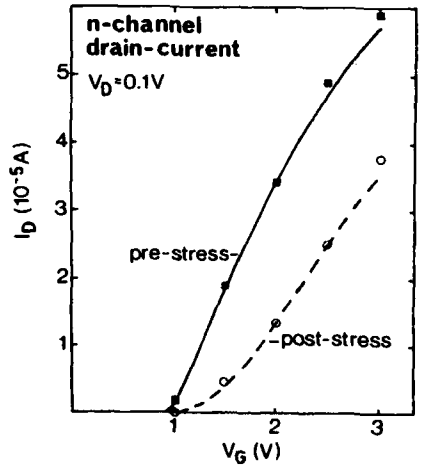


FIGURE 2
Drain-current vs. gate voltage; experiment: solid line (pre-stress), dashed line (post-stress); simulated data: squares (pre-stress), circles (fast interface states)

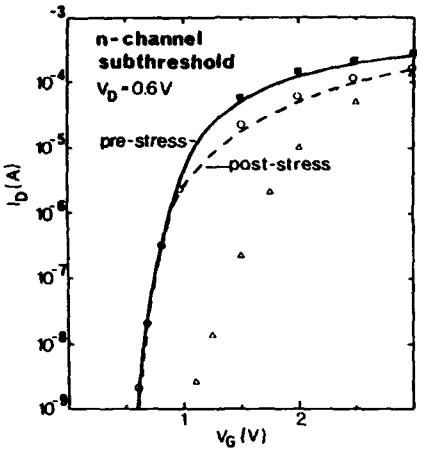


FIGURE 3
Normal mode subthreshold current; symbols defined in figure 2, triangles (simulation with fixed negative oxide charges)

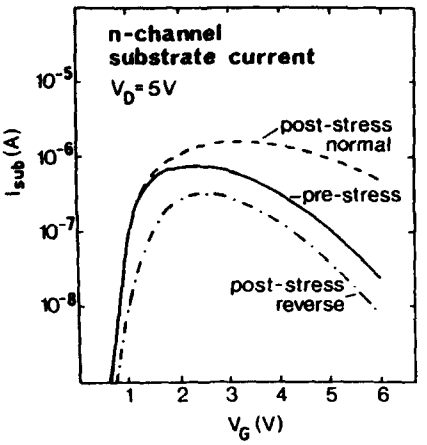


FIGURE 4
Measured substrate current; symbols defined in figure 1

Our investigation takes into account two different effects due to the presence of surface charges in degraded transistors:

i) The potential effect - Surface charges change the electric potential in the adjacent channel region.

ii) The mobility effect - Charges localized at the SiO_2/Si-interface act ás Coulomb centers, thus reducing the channel mobility. To zero order, the approach of Sun and Plummer /2/ is used here. It turns out that for the considered gate length the mobility effect is of little importants.

The filling and the influence of the localized interface states is calculated in our modified version of MINIMOS in a self-consistent manner because the surface potential is strongly influenced by the presence of surface charges. To this end a density of states for the fast interface states is introduced. For the sake of simplicity a linearly increasing function between midgap and the band edges is used.

A first attempt to fit the experimental results simulating fixed negative charges alone shows that negative charges are necessary but their amount has to be adjusted for different gate voltages. This is exactly what is provided by acceptor-type fast surface states. Figures 2, 3, 5, and 6 show the simulated results including fast interface states. All simulated currents are obtained with one parameter set. In comparing Fig. 1 and 5 we see: The simulation delivers the same typical asymmetry between normal and reverse mode as found in experiment. This asymmetry is correlated to a strong localization of the damaged region near drain. Fig. 6 shows the right behavior of the simulated bulk current. Normal and reverse mode substrate currents are shifted into opposite directions. Only n e g a t i v e charges (or negatively charged surface states) are suited to deliver the sign of this shift in simulation. In Fig. 2 and 3 the excellent agreement between experimental results and simulated data using fast interface states is demonstrated, the triangles in Fig. 3 show the failure of the attempt to fit the measurements with fixed oxide charges. For all these calculations the charge distribution is located in a region of 100nm before the drain region. This choice of position and spread gives the right asymmetric behaviour for normal and reverse mode I_D vs. V_D characteristics. The

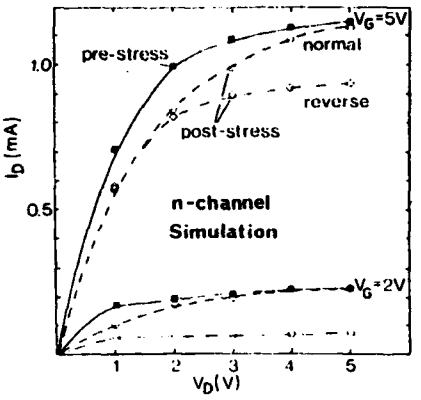


FIGURE 5
Drain current vs. drain voltage - simulation; symbols defined in figure 2, rhombs (post-stress reverse)

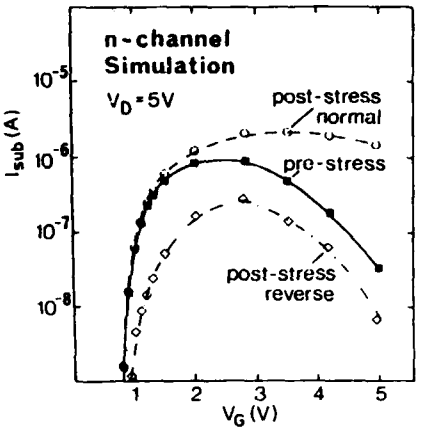


FIGURE 6
Simulated substrate current; symbols defined in figures 1 and 2

overall agreement is achieved without assuming additional fixed charges.

DEGRADATION OF P-CHANNEL MOSFET's

a) Experimental results:

1µm-CMOS p-channel devices with a 0.25µm wide etched spacer were used. The boron source drain implantation was done after the spacer etching.

Figures 7-9 show the shifts in device characteristics after exposition to a stress of V_D=-8V V_G=-2V, which is the voltage condition of maximum degradation, for 15000 seconds. To simplify the discussion only absolute values of currents and voltages are considered in the following.

In p-channel devices drain currents are increased (Fig. 7) and the subthreshold curve is shifted to smaller gate voltages (Fig. 8) after stress. The bulk current in normal mode operation is drastically diminished whereas the reverse mode measurement shows an increase (Fig. 9).

b) Discussion of the results

The overall increase in drain current and the shift of the subthreshold current to smaller gate voltages suggests the existence of negative charges. Thus a simulation was performed assuming fixed negative charges near drain. Amount and position of the oxide charge was fitted to the experimental data. A charge density of 2×10^{12}e/cm^2 within a spread of 100nm delivers good agreement between experimental results and simulation. The strong asymmetry of the bulk current between normal and reverse mode can be understood as follows:

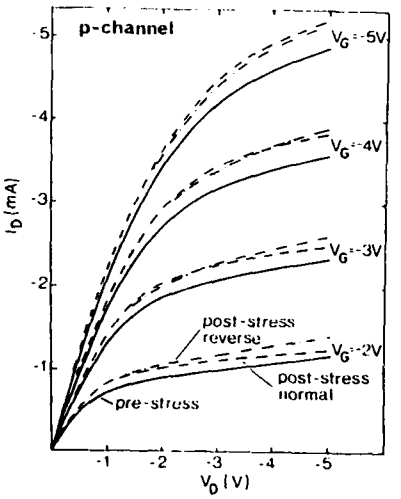


FIGURE 7
Drain current vs. drain voltage - experiment; symbols defined in figure 1; stress conditions: V_G=-2V, V_D=-8V, t_{stress}=15000s

In the reverse mode operation the charges, now near source, reduce the channel length, thus increasing the potential drop at the drain side. This leeds to an elevated bulk current. In normal mode, the oxide charges reduce the electric field peak near drain, impact ionization and henceforth the substrate current are diminished. No influence of fast interface states is found in the p-channel device: The effect of degradation is nearly the same for different (characterization) gate voltages.

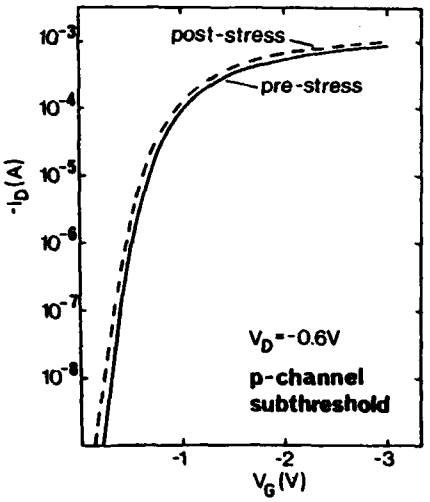


FIGURE 8
Measured subthreshold current; symbols defined in figure 1

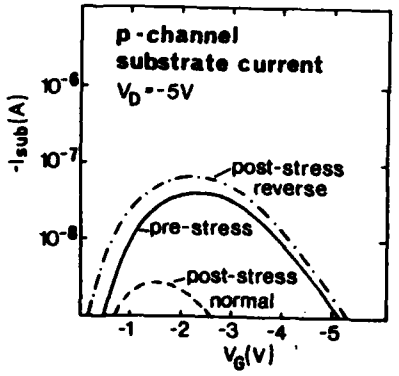


FIGURE 9
Measured substrate current; symbols defined in figure 1

Negative charges in the gate oxide of p-channel transistors are not contradictory. A large amount of electrons is generated by impact ionization at the stress bias. A connection between the injection of these electrons into the oxide and the degradation of the p-channel device should exist. This is backed by the fact that negative gate current and p-channel degradation show a similar gate voltage dependence /3/,/4/.

SUMMARY

We showed a comparison between experimental results and simulation of degraded n- and p-channel MOSFET's. Throughout this investigation we considered conventional devices where the damage takes place near drain under high field stress conditions. In n-channel devices degradation can be explained with acceptor-type fast interface states in the upper half of the Si-band gap. In p-channel devices negative fixed oxide charges can explain the observed data.

A more detailed analysis will be presented in a forthcoming publication /5/.

REFERENCES

/1/ W. Hänsch and S. Selberherr, IEEE Transactions on Electron Devices, ED-34, pp.1074-1078, 1987.

/2/ S. C. Sun and J. D. Plummer, IEEE Transactions on Electron Devices, ED-27, pp.1497-1508, 1980.

/3/ Fig. 1 in K. Kawabuchi, M. Yoshimi, T. Wada, M. Takahashi and K. Numata, IEEE Transactions on Electron Devices, ED-32, pp.1685-1687, 1985.

/4/ Fig. 1a in W. Weber and F. Lau, IEEE Electron Device Letters, EDL-8, pp.208-210, 1987.

/5/ A. Schwerin, W. Hänsch and W. Weber, submitted to IEEE, Trans. Electron Devices.

P2.2.17

Simulations of aging effects in MOS transistors

C. Bergonzoni
SGS Microelettronica
via C.Olivetti 2
20041 Agrate Brianza (MI)-Italy

B. Doyle
BULL
rue Jean-Jaures 78340
Les Clayes sous Bois-France

1 Introduction

The study of the effects of aging on MOS transistors leads to the identification of a wide set of phenomena affecting the electrical characteristics of a stressed device. Degradation of transconductance and threshold voltage are commonly regarded as the most important effects, and several mechanisms have been proposed for their interpretation; the common statement of the models is in any case that hot carrier creation and injection into the gate oxide stays at the origin of any observed aging degradation, while there is not agreement about the way it acts on the physical status of the device after the stress [1,2,3]. The purpose of clarifying the description of the aging effects has been here pursued mainly by means of 2-D device simulations, that allowed the analysis both of stress conditions and stressed devices operation; the comparison of simulations to experimental data shows that the simple location of fixed charge in different regions of the gate oxide is able to reproduce the observed device behaviors, while a further confirmation of the validity of this approach is given by a Fowler-Nordheim tunneling experiment in which charge is injected and trapped exclusively in the overlap region between gate and drain.

2 Simulations

The 2-D device simulator we used during this work is the Poisson and continuity equations solver HFIELDS, developed at the University of Bologna; it was inserted in a simulation environment called IDAS, which allows the set up of the geometric structures needed for the simulations and the extraction of the results by means of various graphic tools. The simulated devices were n-channel and p-channel MOS transistors, with gate length from 1.25 μm to 1.75 μm.

2.1 Simulation of the stress conditions

The identification of the regions in which injection takes place is obviously a major requirement for a description of aging based on the role of trapped charge.

For each sign of the channel three stress situations have been examined:

- A- High drain and high gate
- B- High drain and low gate
- C- Low drain and high gate

This preliminary work has led to the identification of two possible injection regions, in which the sign of trapped charge depends on applied bias; these regions, that will be in the following referred as region 1 and region 2, lie, as shown in fig.1, over the terminal zone of the channel near the drain diffusion and in the gate overlap zone. In particular, simulations show that in the bias situation A- minority carriers can be injected in region 2, in bias case B- majority carriers can be injected in region 1, while in case C- minority carriers can be injected in region 1. The terms 'majority' and 'minority carriers' are here used with regard to the bulk doping type.

Simulations and experiment suggest that, for given high drain voltage, a critical gate voltage defines the boundary between hole and electron injection. It can be seen [4] that hole injection is possible only with gate polarisation much lower than drain voltage; this gives for the gate critical bias a value just above the threshold voltage. It should be pointed out that since our simulation program doesn't provide at present a carrier multiplication model, the identified regions are to be considered as *possible* injection regions, as the simulation doesn't give any information about the energy of the carriers and the actual multiplication rate, but it only reports the intensity and direction of the electric field vector.

2.2 Simulations of stressed devices

Simple box-like charge distributions have been localised at the interface between gate oxide and silicon, following the guide lines coming from previous simulations. Regions 1 and 2 have been given a spatial extent of 0.15 μm beyond and within the lateral drain junction. Three fundamentally different behaviors have been observed, depending on the spatial distribution and on the sign of fixed electric charge: they can be summarised as follows:

- Bias case A-:High gate, high drain; minority carrier injection in region 2: decrease of transconductance.
- Bias case B-:Low gate, high drain; majority carrier injection in region 1: increase in transconductance.
- Bias case C-:High gate, low drain; minority carrier injection in region 1: increase in threshold voltage.

Figures 2,3,4 show the transfer characteristics for an *n-channel transistor* before and after stress in the three situations above summarized.

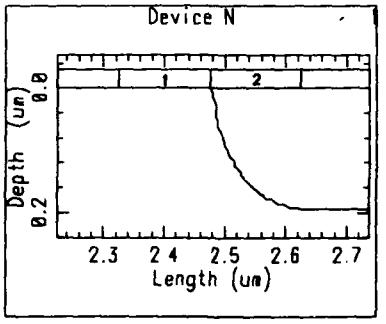


Figure 1: Schematic structure of the simulated devices, with indication of injection regions 1 and 2

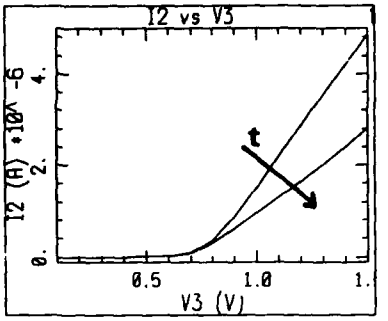


Figure 2: N-channel transistor transfer characteristics before and after stress ; negative fixed charge in region 2 ($4 \cdot 10^{12} cm^{-2}$)

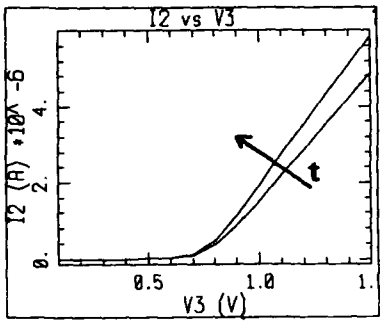


Figure 3: N-channel transistor transfer characteristics before and after stress ; positive fixed charge in region 1 ($4 \cdot 10^{12} cm^{-2}$)

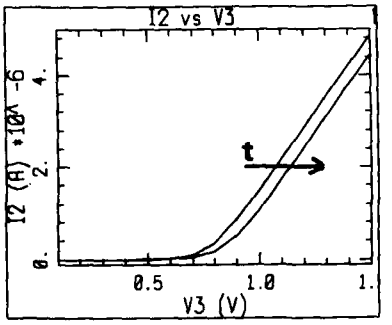


Figure 4: N-channel transistor transfer characteristics before and after stress; negative fixed charge in region 1 ($3 \cdot 10^{11} cm^{-2}$)

3 Experimental

Hot carrier stressing was carried out on 1.5 μm gate length, 25 μm gate width, wet gate oxide transistors. The conditions of stressing were $V_g = 5V, V_d = 9V$. The $I_d - V_g$ curves were taken after fixed times on a log scale, with interval $\Delta(log(t)) = 0.1$. It can be seen (fig.5) that the stressing results in a degradation of g_m, the transconductance, but that the degradations are limited to gate voltages above the threshold voltage - that is, there is no shift in threshold voltage. Analogous experiments, whose results are not pictured here owing to space limitations, have been performed under different stress conditions and on p-channel devices, showing the behavior predicted by simulations. In order to verify the localization of

the trapped charge, given by the above simulations, Fowler-Nordheim (F-N) injection measurements were also performed. Here, electrons are injected from the gate to the drain and the effect of this injection on the $I_{ds} - V_{gs}$ characteristics is examined. The injection from the gate is performed in the constant current mode, and in order to localize it to the gate-drain overlap region, a voltage is applied to the drain (10 V) with the source grounded. The injection thus takes place solely between the gate (biased at -20 V) and the drain, as the potential difference is here of the order of 30 Volts. In this configuration, the field between gate and source, or gate and channel, is too small to result in F-N injection.

Figure 6 shows the result of this stress injection on the $I_d - V_g$ characteristics. Curve a) shows the characteristics before injection. Curve b) was obtained by injecting for 10 seconds with a gate current of 50pA. It can be seen that the effect is to cause a change in the transconductance, but at gate values above the threshold voltage. This is identical to the simulation of positive charge in figure 3 and indicates that, initially, holes are trapped in the oxide. For longer times (500 and 1000 seconds, curves c) and d) respectively) and higher injection levels (500 pA) the $I_{ds} - V_{gs}$ characteristics can be seen to degrade, but once again only above threshold voltages. Since in this experiment we control the spatial distribution of injected charge in the overlap region, we can compare these transfer characteristics to the simulations shown in fig.2, concluding that charge trapping above the drain in the gate oxide confirms experimentally the simulation results for an analogous charge distribution, that are indeed seen in *real* stress conditions.

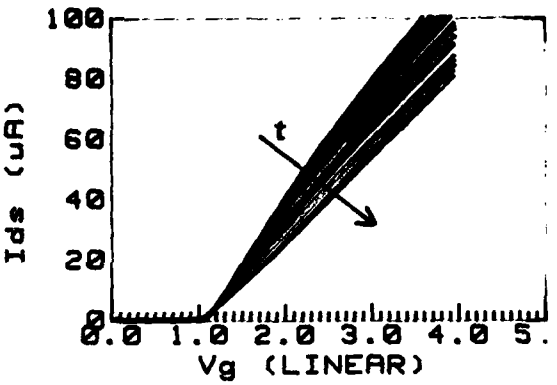


Figure 5: Experimental transconductance degradation in stressed n-channel transistors

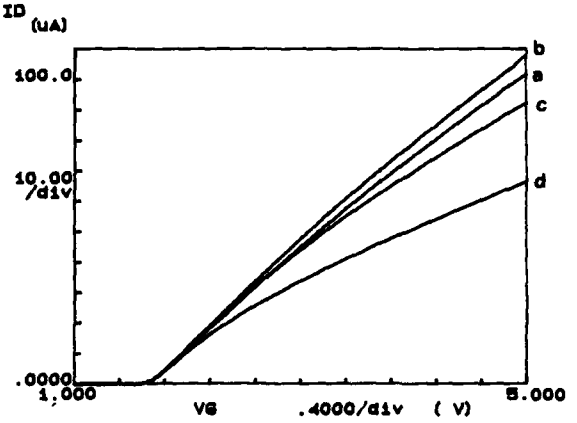


Figure 6: Degradation induced by a Fowler-Nordheim tunneling experiment, with charge injection in the gate-drain overlap region. a) unstressed device, b) 10 seconds-50 pA, c) 500 seconds-500pA, d) 1000 seconds-500pA

4 Discussion and conclusions

From a qualitative physical point of view, the decrease of transconductance can be interpreted as the effect of a thin potential barrier near the drain junction, which limits the number of carriers reaching the drain itself (fig.7); the increase of transconductance is a simple effect of channel shortening, due to the strong inversion of a region of the channel, that operates like a kind of drain extension (fig.8); finally, a fixed minority charge localization over an extended region of the channel will induce a *channel-stop* zone, preventing carriers from getting to the drain until the local threshold voltage of the stressed region has been reached and even that part of the channel has been inverted (fig.9).
It will be shown elsewhere [5] that it is possible to state a semi-empirical 1-D analytic model, giving account for these and other aging effects, such as transconductance overshoots and distortions.

The simulation of localised layers of fixed charge in different regions of the gate oxide seems so to be able to reproduce several experimented aging effects; it should nevertheless be pointed out that this doesn't exclude the existence of other and different mechanisms, such as mobility degradation, causing modifi-

cations in the electrical characteristics of a stressed device. What here is shown is that an analysis of the simpler effect of stressing, that is hot carrier injection and trapping in the gate oxide, is a valid tool for the identification and prevision of the kind of degradation that a device will undergo if subjected to certain stress conditions. The existence and effect of stress induced

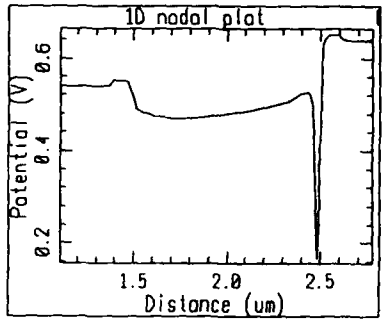


Figure 7: Potential barrier for minority carrier injection in region 2,resulting in transconductance decrease (see fig.2)

fast interface states are presently under experimental and theoretical analysis, and significant elements have been collected in this field. The study of the electric fields into the device can also give significant informations about the influence of device geometry and structure (oxidation-induced gate bending, LDD dose and extension, etc.) on the intensity of hot carrier creation and injection and, as a consequence, on the importance of the stress effects.

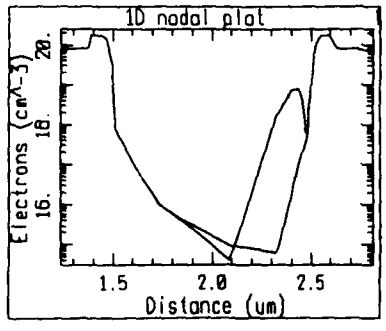


Figure 8: Effect of majority carrier injection in region 1 on minority carrier concentration in the channel, resulting in transconductance increase (see fig.3)

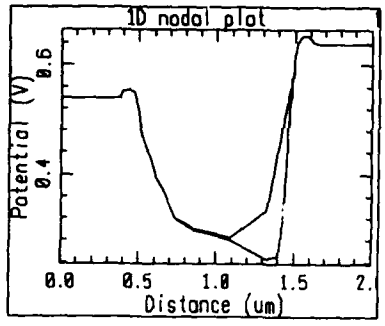


Figure 9: Potential barrier for minority carrier injection in region 1, resulting in threshold voltage increase (see fig.4)

Aknowledgments

This work has been supported by ESPRIT Project 554-SPECTRE. The authors wish to thank ing. Claudio Lombardi for helpful discussions and simulation support.

References

[1] R.B.Fair and R.C.Sun, IEEE Trans. El. Dev., Vol. ED-28, p.83, 1981

[2] H.Gesh, J.P. Leburton and G. Dorda,IEEE Trans. El. Dev., Vol. ED-29, p.913, 1982

[3] T.Tsuchiya, T.Kobayashi and S.Nakajima, IEEE Trans. El. Dev., Vol. ED-34, p.586, 1987

[4] M.Tosi, L.Baldi and F.Maggioni, 'The role of holes and electrons in the aging of MOS transistors', this volume

[5] C.Bergonzoni and P.Caprara, not yet published

P2.2.18

GAMMA-RADIATION EFFECTS IN CMOS TRANSISTORS

S. Golubović, S. Dimitrijev, D. Župac, M. Pejović, N. Stojadinović

Faculty of Electronic Engineering, University of Niš
Beogradska 14, 18000 Niš, Yugoslavia

Effects of gamma-radiation in Al-gate and Si-gate CMOS transistors have been investigated in this paper. It has been found that, besides the well-known threshold voltage instabilities, radiation causes the gain factor instabilities as well. These instabilities of CMOS transistor electrical parameters have been found to be caused by significant increase of positive gate oxide charge density and somewhat smaller increase of interface trap density. While the positive gate oxide charge increase has been explained in terms of the broken bond model, a new model has been proposed to elucidate the interface trap increase as well as existing partial compensation of the created positive gate oxide charge. Finally, note that no significant difference in radiation hardness between Al-gate and Si-gate CMOS transistors has been observed.

1. INTRODUCTION

The influence of radiation on MOS devices has been extensively investigated, and as emphasized in the textbook by Nicollian and Brews [1], the main radiation effects are due to the creation of both positive gate oxide charge and interface traps. However, as the results about the types and quantities of the created gate oxide charge and interface traps were generally obtained by analysis of C-V curves of MOS capacitors, there was no relationship to degradation of electrical characteristics of CMOS transistors. On the other hand, in the papers where CMOS transistors have been used to investigate instabilities of their electrical characteristics, results concerning analysis of instability mechanisms (changes in the gate oxide charge and interface traps) remained in the scope of qualitative descriptions.

In this paper, an investigation of gamma-radiation effects in both Al-gate and Si-gate CMOS transistors is made using the recently proposed method [2] for calculation of the gate oxide charge and interface trap densities by MOS transistor transfer characteristics. The obtained results are analyzed to provide a structural and/or chemical insight into the problem concerning creation of the oxide charge and interface traps.

2. EXPERIMENTAL RESULTS

Al-gate CMOS transistors used were from CD4007UB integrated circuits, produced by the standard technological procedure for CD4000 series CMOS integrated circuit manufacturing, while Si-gate CMOS transistors were from specially designed test chips RT20AC, produced by the standard technological procedure for 18000 series CMOS microprocessor manufacturing. It should be emphasized that the gate oxide (90 nm thickness) was formed in a wet oxygen with presence of 0.2% $C_2HC]_3$, and the chips were assembled in ceramic dual-in-line packages.

The transistors under investigation were irradiated using the ^{60}Co radiation source, with no bias applied. After given radiation doses, the CMOS transistor transfer characteristics were measured in order to determine the threshold voltage and gain factor which are the most important electrical parameters of a MOS transistor.

The threshold voltage behaviour of both Al-gate and Si-gate CMOS transistors during the irradiation is shown in Fig. 1a. As can be seen, the threshold voltage of PMOS transistors* increases in the case of either gate-type. On the other hand, the threshold voltage of Si-gate NMOS transistors decreases all the way, while the threshold voltage of Al-gate NMOS transistors decreases only up to the dose of about 10^5 rads,

* The absolute values of the threshold voltage of PMOS transistors are considered in this paper.

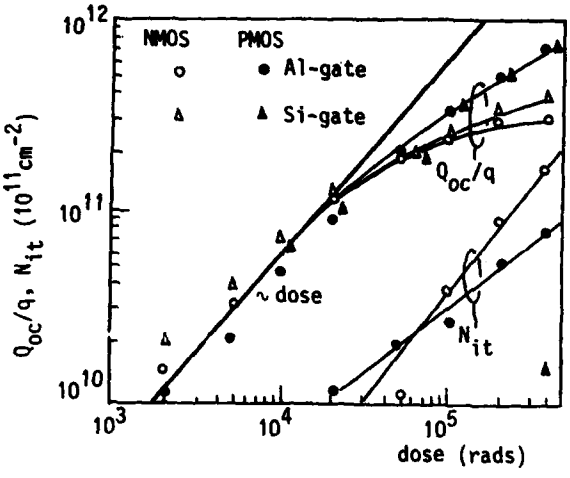


FIGURE 2
Effects of gamma-radiation on gate oxide charge and interface traps of CMOS transistors

transistor electrical parameters are caused by significant increase of the positive gate oxide charge in both Al-gate and Si-gate transistors, and somewhat smaller increase of the density of interface traps in Al-gate transistors.

The fact that the radiation creates a significant amount of the positive gate oxide charge is well known [1] and has been explained in terms of broken bond model [6]. Namely, because of the great flexibility of the Si-O network and Si-O bonds, a Si-O bond can break by rotation and redistribution of the bonding electrons, without significant atomic displacement, which can hardly be caused by ionizing radiation anyway [7]. During that process, electrons are removed from the atomic bond [6] and, being free, drift away from the oxide to a more positive electrode (usualy, it is the gate), while holes remain captured at either the formed trivalent-silicon traps ($\equiv\overline{Si}^{*}$) [7,8] or nonbridging oxygen traps ($\equiv\overline{Si\text{-}O}^{\cdot}$) [6,8]. These holes are trapped very close to the oxide valence band (at about 0.4 eV above the valence band [7]) and can easily move towards the oxide-silicon interface [9] where much deeper hole trap levels are located [8,10].

* The bars indicate that the Si atom is bonded to three O atoms, and the dot represents the unsaturated bond.

A linear dependence of the positive gate oxide charge increase on the radiation dose should be expected if the charge was created by the mechanism described above. As can be seen from Fig. 2 we have found such a behaviour up to the dose of about 2×10^{4} rads. In addition, significant differences in the densities of the created positive gate oxide charge between NMOS and PMOS transistors, as well as between Al-gate and Si-gate transistors, have not been observed in this region, what is also quite sensible.

However, at higher doses the rate of increase of the positive gate oxide charge is reduced, causing a departure from initial linear dependence of the oxide charge on the dose (Fig. 2). It has been belived [1,11] that it is due to partial compensation of the positive gate oxide charge caused by electron tunneling from the silicon which is enhanced by the electric field of the trapped positive gate oxide charge. However, in our opinion this mechanism cannot be responsible for this effect because of at least three reasons. First, the significant difference in the density of net positive gate oxide charge between NMOS and PMOS transistors, which can clearly be seen in Fig. 2, could not be explained by this mechanism. Second, the gate oxide electric field, caused by the work function difference and the gate oxide charge is too low (about 0.1 MV/cm) to be assumed as a posible cause of electron tunneling. Third, tunneling of electrons from the silicon into the oxide electron traps has not been found even after stressing by high positive gate oxide electric fields [2].

Instead of electron tunneling, the partial compensation of the created positive gate oxide charge can more duly be explained by the following mechanism. First, remember that holes created by irradiation move towards the oxide-silicon interface. As the large amount of rather weak Si-H and Si-OH bonds [7] (formed by the anneaing of interface traps, which is commonly used during device production [1]) exists at the oxide-silicon interface, the holes reaching the

starting to increase beyond that dose. It should be noted that the presented behaviour of the threshold voltage of Al-gate NMOS and PMOS transistors has also been observed in many experiments performed by other investigators, and it has been generally established as a typical behaviour [1]. On the contrary, investigations for Si-gate transistors have not been made so widely, and published results [3], expressing that the threshold voltage shifts of Si-gate transistors are much greater comparing to Al-gate transistors, differ from those given in Fig. 1a.

The gain factor behaviour of the investigated transistors is shown in Fig. 1b. It can be

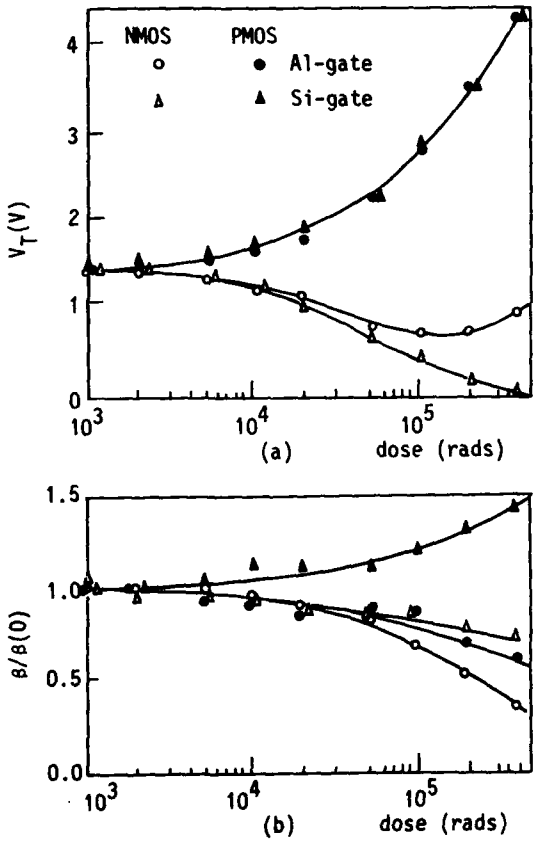


FIGURE 1
Effects of gamma-radiation on threshold voltage (a) and gain factor (b) of CMOS transistors

seen that the radiation leads to the gain factor decrease in Al-gate NMOS, Al-gate PMOS, and Si-gate NMOS transistors, as well as to the gain factor increase in Si-gate PMOS transistors. As regards Al-gate transistors, it is in accordance with previously published results [4,5] implying the fact that radiation reduces mobility of channel carriers, i.e. the gain factor. However, to our knowledge, there are no published results concerning the behaviour of either the gain factor, i.e. mobility of channel carriers, or transconductance of Si-gate transistors.

3. ANALYSIS

It is very important to determine and analyze instabilities of the gate oxide charge and interface traps which cause the changes in the threshold voltage and gain factor described above. It can be performed on the basis of the threshold voltage and gain factor data according to the recently proposed method [2]. In this method, densities of the gate oxide charge and interface traps of CMOS transistors can be determined using the models for their influence on the threshold voltage and gain factor. These models for NMOS and PMOS transistors are given by the following expressions [2]:

$$V_T = V_{TO} \mp \frac{Q_{oc}}{C_{ox}} + \frac{qN_{it}}{C_{ox}} \qquad (1)$$

$$\beta = \frac{\beta_o}{1 \pm \alpha_{oc} Q_{oc}/q + \alpha_{it} N_{it}} \qquad (2)$$

In expressions (1) i (2) V_T and β are the threshold voltage and gain factor, respectively, which account for effects of the gate oxide charge (Q_{oc}) and interface traps (N_{it}); V_{TO} and β_o are the threshold voltage and gain factor, respectively, when there are no gate oxide charge and interface traps; α_{oc} and α_{it} are the coefficients in the model for gain factor expressing the influence of the gate oxide charge and interface traps, respectively ($\alpha_{oc}=0.2\times10^{-11}cm^2$ and $\alpha_{it}=3.5\times10^{-11}cm^2$ [2]).

The changes in the gate oxide charge and interface traps of Al-gate and Si-gate CMOS transistors during the irradiation, as obtained using expressions (1) and (2), are shown in Fig.2.

Obviously, the observed instabilities of CMOS

interface give rise to their dissociation by the following reactions:

$$Si\text{-}H + h^+ \rightarrow Si^+ + H^\cdot \quad (3)$$

$$Si\text{-}OH + h^+ \rightarrow Si^+ + OH^\cdot \quad (4)$$

The released hydrogen atoms and hydroxil groups easily diffuse through the oxide [7], after which the trivalent-silicon defects formed become stable. It is well known that those interface trivalent-silicon atoms introduce the allowed energy levels into the forbiden gap of the silicon [1], and what is important, a part of those atoms (having energy levels below the Fermi level) are neutralized by free silicon electrons. In this way, a part of positive gate oxide charge is compensated. Note that this compensation is more pronounced in NMOS transistors, what results from the higher position of the Fermi level under strong inversion condition.

On the other hand, the trivalent-silicon atoms having energy levels within about 0.1 eV above and below Fermi level, continously capture and release free carriers, which mechanism reduces mobility [2] (expression (2)). Such trivalent-silicon atoms appear as active interface traps and their density, corresponding to strong inversion condition, is measured by the method used. It can be seen from Fig. 2 that these interface traps are more pronounced in Al-gate transistors (esspecialy in NMOS transistors) and almost negligible in Si-gate transistors. This advantage of Si-gate transistors is a consequence of manifold high-temperature treatment of the gate oxide, inherent in self-aligned, Si-gate technology.

4. CONCLUSIONS

The gamma-radiation effects in Al-gate and Si-gate CMOS transistors have been investigated in this paper. It should be emphasized that we have obtained typical, and in literature widely reported threshold voltage shifts of the irradiated CMOS transistors, but additionally, we have presented the accompanying gain factor instabilities of those transistors. Note that rather similar radiation responses of both Al-gate and Si-gate CMOS transistors have been obtained indicating that there is no significant difference in radiation hardness between the two technologies.

The underlying creation of the gate oxide charge and interface traps has been analyzed using the recently proposed method [2] for the calculation of the oxide charge and interface trap densities from the threshold voltage and gain factor data. The positive gate oxide charge density increase, caused by breakage of ≡Si-O-Si≡ bonds, as well as interface trap density increase, caused by breakage of Si-H and/or Si-OH bonds (which mechanism is, in our opinion, also responsible for the partial compensation of the created positive gate oxide charge), have been found.

REFERENCES

[1] Nicollian, E.H. and Brews, J.R., MOS Physics and Technology (John Wiley, New York, 1982).

[2] Dimitrijev, S. and Stojadinović, N., Solid-State Electronics, in print.

[3] Nordstrom, T.V. and Gibbon, C.F., IEEE Trans. Nucl. Science NS-28 (1981) 4349

[4] Burghard, R.A. and Gwyn, C.W., IEEE Trans. Nucl. Science NS-20 (1973) 300

[5] Bellaouar, A., Sarrabayrouse, G. and Rossel, P., IEE Proceedings 132 (1985) 184

[6] Gwyn, C.W., J. Appl. Phys. 40 (1969) 4866

[7] Revesz, A.G., IEEE Trans. Nucl. Science NS-24 (1977) 2102

[8] Sah, C.T., IEEE Trans. Nucl. Science NS-23 (1976) 1563

[9] Hughes, R.C., Ear Nisse, E.P. and Stein, H.J., IEEE Trans. Nucl. Science NS-22 (1975) 2227

[10] Oldham, T.R., Lelis, A.J. and McLean, F.B., IEEE Trans. Nucl. Science NS-33 (1986) 1203

[11] Boesch, H.E., McLean, F.B., Benedetto, J. M. and McGarrity, J.M., IEEE Trans. Nucl. Science NS-33 (1986) 1191

TEMPERATURE INCREASE BY SELF-HEATING IN VLSI CMOS

D.Takacs,J.Trager

Siemens AG,Corporate Research and Development,Microelectronics
Otto-Hahn-Ring 6, D-8000 Munich 83, FRG

Experimental and theoretical investigations were carried out to study the steady-state and transient self-heating effects in VLSI MOST's.The drain and substrate currents and the thermoelectric power in the substrate were used to monitor the internal temperature of the MOST.The results show that the devices operate in nonisothermal conditions due to the temperature rise in the device caused by self-heating in the steady-state and the transient regimes.

1.INTRODUCTION

The self-heating effect caused by power dissipation at high power densities is well known for power MOSFET's [1], for high voltage integrated MOST's [2] and for MOS transistors operated at cryogenic temperatures [3].
The thermal response to an ESD pulse in CMOS input protection circuits was investigated theoretically as well [4].

However,the self-heating effect has not been analysed for VLSI MOSFET's at room temperature. Although maesurements have been carried out for MOST's in the temperature range of 77-300 K [5],no self-heating effect was observed, since the dissipated power density was too small to detect the effect.

As the feature sizes in VLSI CMOS are continously scaled down to submicron dimensions,the power density strongly increases at constant operating voltages. Most of the power in MOS transistors is dissipated in the pinch-off region at the drain side. If the power density is high enough,significant heat is produced in the MOST.This leads to a local temperature increase that changes the output characteristics of the device itself and can thermally influence other devices in its vicinity,leading to thermal coupling between adjacent devices.One consequence of the local temperature rise in MOSTs is a reduced drain current.

In this paper,experimental and theoretical results are presented for the steady-state and transient self-heating effects in VLSI-MOST's at room temperature. The drain and the substrate currents and the thermoelectric power in the substrate were used as monitors of the internal temperature in the MOST. The self-heating effect was also studied by calculations of the steady-state and transient temperature distributions in the substrate.

2. STEADY-STATE SELF-HEATING EFFECT

The cross sectional view of the test structure used for heating by an adjacent MOS transistor is shown in Fig.1. This is a gated inverter structure, consisting of an n- and a p-channel transistor. Investigating one of the transistors, the other one can be used to additionally dissipate power (in the following called transistor heating).

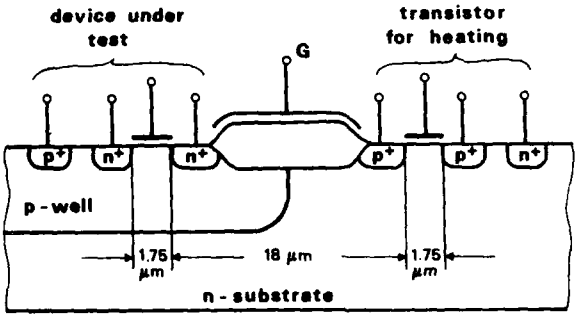


Fig.1: Cross sectional view of the test structure with the device under test and the heating transistor. (Wn = 50um, Wp = 100um).

The influence of the additionally dissipated power of 330mW in the p-channel transistor on the output characteristics of the adjacent n-MOS transistor is shown in Fig.2. The influence of the transistor heating is clearly seen:
with transistor heating,the temperature at the n-MOS transistor increases and the drain current is reduced.
Further experiments show that the drain current reduction depends linearly on the power dissipated in the transistor used for heating.
The same drain current reduction can be achieved without transistor heating by increasing only the ambient temperature Ta=To+ΔT. In this way an equivalent tem-

perature increase ΔTeq can be found for every power P dissipated by the heating transistor. At an elevated ambient temperature Ta=To+ΔTeq without transistor heating and at ambient temperature To with transistor heating, the two output characteristics coincide.

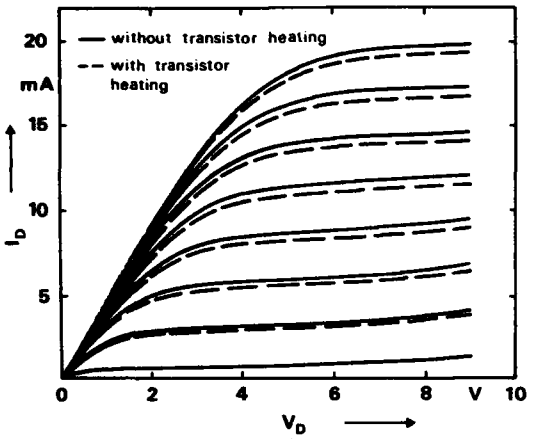


Fig.2: Output characteristics of an n-channel transistor without and with transistor heating at the p-channel transistor (P=330mW).

An equivalent temperature increase ΔTeq = 13 K was found for a dissipated power P=330mW for the geometry given in Fig.1.The spacing between the n-channel transistor and the transistor used for heating is relativly large (about 20um). It is evident that the temperature increase in the heating transistor itself is much higher. The measured drain current reduction and the corresponding equivalent temperature increase caused by the transistor heating are significant considering the relatively large distance.

Fig.3 demonstrates the influence of the transistor heating on the maximum substrate current of the n-MOST at a given operating point. The transistor heating, as expected, reduces the substrate current,being carrier multiplication induced.Increasing the ambient temperature, the substrate current also decreases.

A further evidence for the self-heating effect in VLSI MOS transistors is presented in Fig.4. The thermoelectric power Vthermo as a function of the dissipated power P is shown for different gate voltages of the heating transistor.The corresponding drain currents are also shown.It can be seen, that the thermoelectric power depends only on the dissipated power and is independent of the partial quantities Id or Vds.The same result is obtained if the power is dissipated in a poly-silicon resistor

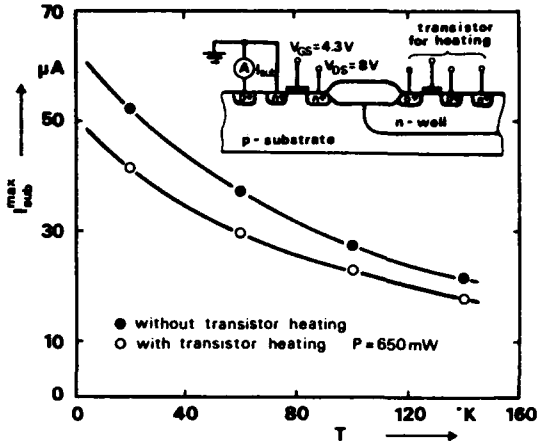


Fig.3: Influence of the transistor heating and the ambient temperature on the maximum substrate current of the nMOST.

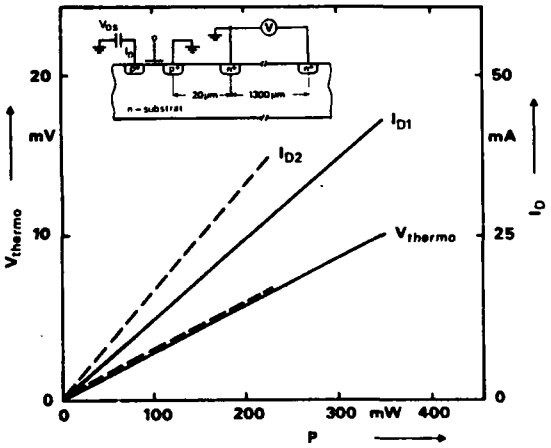


Fig.4: Steady-state thermoelectric power Vthermo vs. power dissipated in the p-channel transistor.

that is totally isolated from the substrate.

The internal temperature in the heating transistor itself can not be determined by electrical measurements in the steady-state. One can only estimate the temperature distribution in the device caused by a steady-state power density, if one solves the heat conduction equation for a given geometry.

The 3D heat conduction equation was solved analytically for the steady-state in a similar manner as described in [6]. The results of the steady-state calculation of the temperature distribution are shown in Fig.5.

A dissipated power of 200mW on the surface area of 5um*5um results in a maximum temperature increase of 130 K. The temperature gradient is very large.

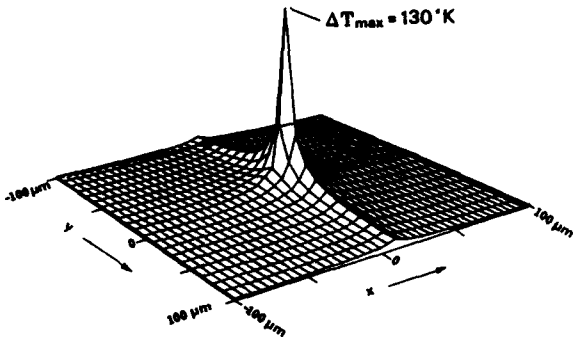


Fig.5: Calculated temperature increase in steady-state for a power of 200mW dissipated in a 5um*5um area on the surface of a chip sized 2mm*2mm.

The measured temperature increase of 10-15 K in Fig.2 and Fig.3 is in good agreement with the results of the rough theoretical estimation for comparable power densities at a distance of about 20um from the heating-source.Fig.5 also suggests that the temperature increase in the heating transistor itself is much higher.

3.TRANSIENT SELF-HEATING

Since integrated circuits usually operate in a dynamic mode, it is important to investigate the time dependence of the self-heating effect.One possibility to study the transient self-heating effect is to measure the time dependence of the drain current reduction.

Single pulse drain current measurements were carried out with pulse rise times in the ns range (typically 10 ns) and pulse durations from 100ns up to a few ms to reach steady-state conditions.

The drain current reduction vs. time is shown in Fig.6 for two different dissipated powers P,(P=Vds*Id in steady-state,recombination processes were neglected). At first,the drain current decreases strongly with time and saturates in the us range.The drain current reduction and thus the intrinsic temperature rise is very fast and most of the effect seems to happen in the first 30-50 ns after power dissipation has started.However, reliable measurements were possible only at times longer than about 30ns.For shorter times,only simulations could be used to conclude what happens at the beginning of the power dissipation.

The earliest measurable drain current reduction for 32mW is about 7%, which probably underestimates the self-heating effect.Lower dissipation leads to lower temperature increase, but the time constant seems to remain the same.

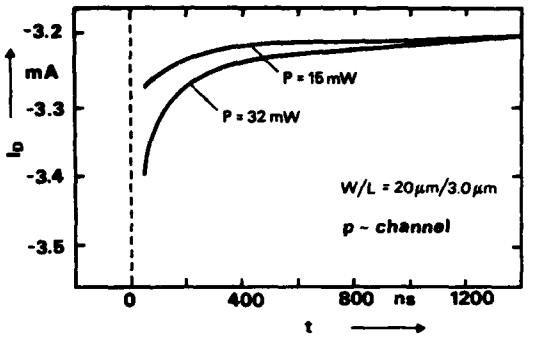


Fig.6: Drain current reduction vs.time for two different dissipated powers in a p-MOST. Power dissipation started at t=0.(Vgs=Vds)

The transient drain current reduction by self-heating as a function of the dissipated power is demonstrated in Fig.7. The drain current reduction is already significant for dissipated powers above 10mW.The slope of the curve is 4.5uA/mW.

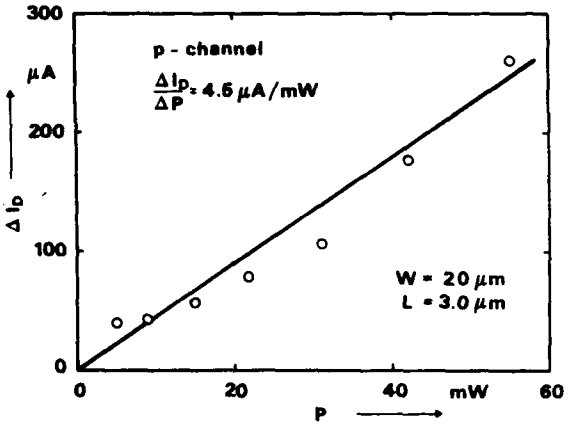


Fig.7: Earliest measurable transient drain current reduction of a p-MOST vs. power dissipated in it.

As device dimensions are scaled down to submicron levels,the problem will be much more severe because the dissipated power density in the MOST strongly increases.Furthermore,the thermal conductivity of the silicon substrate decreases because of the higher doping concentration used to avoid short channel effects in VLSI.Therefore, the channel length dependence of the self-heating has to be investigated.

Fig.8 shows the drain current reduction as a function of time for different channel lengths at constant dissipated power P=32mW.

Decreasing the channel length, the detectable drain current reduction and the time constant decreases. For a channel length of 1um, no drain current reduction can be observed. We conclude from the shorter time constants for shorter channel lengths, that the self-heating process for L=1um already has completed in the first 30ns.

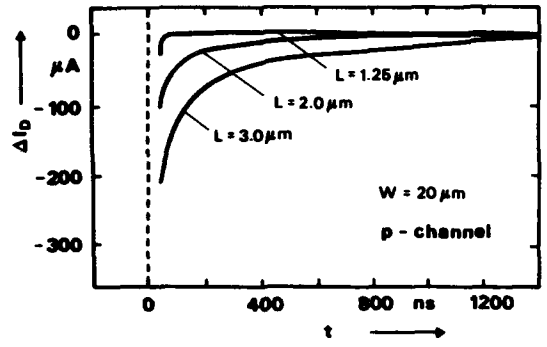


Fig.8: Drain current reduction vs. time for different channel lengths.
The dissipated power in the p-MOST was kept constant at 32mW.

To estimate the self-heating in this time region the time-dependent 2-D heat conduction equation has been solved (as described in [7]).
The homogeneously doped silicon chip has infinite dimensions in this model and the heat-source term has two components: a delta-function distribution adjacent to the drain and an even distribution for the channel area. No heat transfer from the surface is considered.

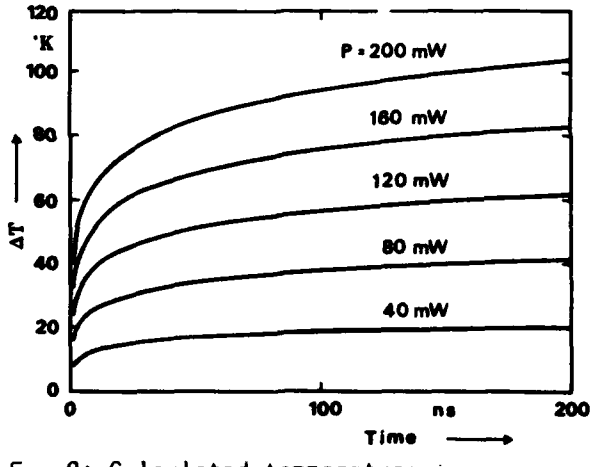


Fig.9: Calculated temperature increase in transient regime vs. time for different dissipated powers in a MOST. (W/L=20um/1.25um)

Fig.9 shows the calculated temperature increase for different dissipated powers in the first 200ns.
In the calculations, one half of the power is dissipated in the channel area and the other one in the pinch-off region (delta-function).
A significant temperature increase takes place in the first few ns, as the results of the experiments indicate.

Furthermore, the self-heating effects in n- and p-MOST in triode or saturation regions seem to be identical.

The investigated nMOST show similar self-heating effects independent of the drain fabrication process (conventional or LDD). However, the transient drain current reduction is more complicated: the nMOST with short channel lengths can operate in bipolar mode with positive temperature coefficient of the corresponding collector current.

4. CONCLUSIONS

Steady-state and transient self-heating effects were investigated in VLSI MOST at 300 K. A consequence of the self-heating is a reduced drain current. The drain current reduction will be even more severe at higher power densities when device features will be further scaled down in the future.

The transient drain current reduction is faster for shorter channel lengths. Thus most of the heating effect seems to happen in the first few ns for L<=1um.

These results suggest that the MOSTs in VLSI operate in nonisothermal conditions.

5. REFERENCES

[1] P.S.Barlow, R.G.Davis, M.J.Lazarus, C. Eng: 'Negative Differential Output Conductance of Self Heated Power MOSFETs' IEE PROC., Vol 133, Pt.I, No.5, OCT. 1986

[2] D.Sharma, J.Gautier, G.Merckel: 'Negative Dynamic Resistance in MOS Devices', IEEE J.of SSC, Vol SC-13, No.3.78

[3] D.P.Foty, S.L.Titcomb: 'Thermal Effects in n-Channel Enhancement MOSFET's Operated at Cryogenic Temperatures' IEEE ED-34, No.1, Jan.1987, pp.107-113

[4] G.Krieger: 'Thermal Response of Integrated Circuit Input Devices to an Electrostatic Energy Pulse' IEEE Trans on ED, Vol.ED-34, NO.4, 1987

[5] S.S.Sensic and G.R.Craig: 'Thermal Effects in JFET and MOSFET Devices at Cryogenic Temperatures', IEEE Trans.ED, Vol ED-19, NO.8, p.933, 1972

[6] H.S.Carslaw and J.C.Jaeger: 'Conduction of Heat in Solids', 1976

[7] D.K.Sharma and K.V.Ramanathan: 'Modeling Thermal Effects on MOS I-V Characteristics', IEEE EDL-4, NO.10, 1983

P2.2.20

CHARGE LIMITED BREAKDOWN IN MOS CAPACITORS

E.A. AMERASEKERA and D.S. CAMPBELL

Department of Electronic and Electrical Engineering,
Electronic Component Technology Group,
University of Technology, Loughborough, Leicestershire, LE11 3TU, U.K.

Investigations into the electrical breakdown of p-Si MOS capacitor structures have been conducted for both pulsed and continuous voltage stress conditions. The experimental results show that the threshold voltages for breakdown are dependent on the amount of charge supplied to the MOS system. An analytical treatment of the breakdown conditions with both positive and negative polarity pulses has been given. The breakdown under continuous stress conditions up to 100 ms has been shown to be related to the duration of the applied voltage and, hence, the charge injection required to provide sufficient energy to initiate damage.

1. INTRODUCTION

Recent investigations into the electrical breakdown of SiO_2 under high field stress conditions, have indicated that the sensitivity of MOS structures is related to the charge injection and conduction mechanisms in the dielectric [1] [2] [3]. Other workers [4] [5] [6] have suggested that breakdown is influenced by the applied electric field. Most studies to date have concentrated on the breakdown in constant electric fields. Maximum breakdown thresholds in these conditions are in the region 5-10 $MVcm^{-1}$.

Previous work by the authors [7] [8] have shown that MOS capacitors can withstand fields up to 50 $MVcm^{-1}$ under pulsed conditions. Such fields are commonly experienced in Electrostatic Discharge (ESD) damage and in some cases when large voltage transients are obtained in the course of normal operation. The investigations presented in this paper examine the mechanisms of pulsed breakdown and the influence of charge carriers on the breakdown process.

Experimental results showing the relationship between amount of charge and the breakdown voltage (V_{bd}) under both pulsed and constant voltage conditions are given. A comparison is then made of the experimental data with a theoretical treatment of the charge conduction process, in order to develop a model of the breakdown process with both positive and negative pulses.

2. EXPERIMENTS

2.1 The Device Structures

p-Si MOS capacitors were used in the experiments. They were supplied as part of MOS process assessment wafers by Plessey Research. Figure 1 illustrates schematically a cross-section of the device.

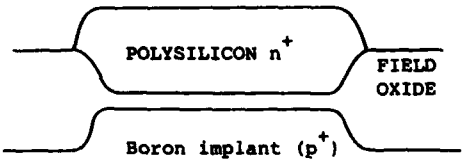


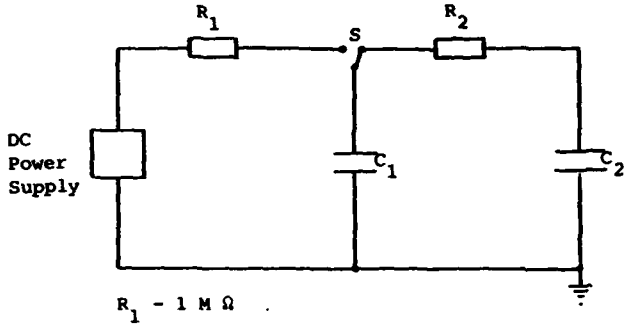
Figure 1 Schematic cross-section of MOS capacitor structures.

A n^+-polysilicon gate was deposited on a thermally grown SiO_2 layer 400Å thick. The p-type Si substrate had a dopant concentration of approximately $4 \times 10^{15} cm^{-3}$. The capacitors

had an area of 49 x 10^{-5} cms^2. Preferential boron doping (by implantation) of the Si surface to a concentration of 1 x 10^{16} cm^{-3} was also present in some capacitors, thereby simulating the gate-oxide capacitors in standard NMOS transistors. Both implanted and unimplanted devices were used in the experiments.

2.2 The Voltage Pulse Generator

The voltage pulse was generated using the discharge system described in Figure 2.



R_1 - 1 M Ω
R_2 - 1.5 k Ω
C_1 - DISCHARGE CAPACITOR
C_2 - MOS CAPACITOR

Figure 2 Circuit used for voltage pulse generator

This circuit is based on the MIL-STD-883C model for testing Electrostatic Discharge (ESD) sensitivities of devices. A commercially developed system, the Hartley AUTOZAP, capable of producing voltage pulses of varying magnitudes with rise times of the order of 100 ns was used in the experiments. A typical voltage pulse is shown in Figure 3.

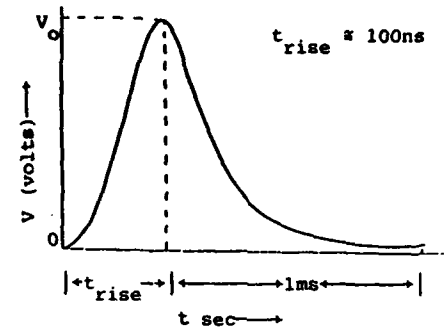


Figure 3 Voltage pulse wafeform

2.3 Measuring Techniques

The static discharge system also measured the current-voltage characteristics of the capacitor before and after application of the pulse. A predefined change in leakage current (taken as ±2μA) at a given gate voltage, after application of the pulse, was taken to indicate breakdown.

Continuous voltage stress was provided by a KEITHLEY 230 programmable voltage source. The system was connected as shown in Figure 4.

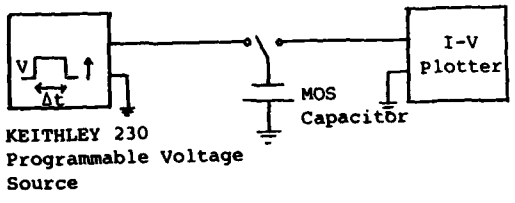


Figure 4 Circuit used for continuous voltage stress experiments

Breakdown was monitored using the current-voltage measuring system available on the Hartley Autozap. This ensured compatibililty with results obtained for pulsed breakdown.

2.4 Pulsed Voltage Breakdown Thresholds

The threshold voltages for breakdown under pulsed conditions were determined by varying the discharge capacitor, C_1, as shown in Figure

2. The value of the capacitor used in the MIL-STD model was 100pF, and the breakdown voltage using this discharge capacitor is usually taken as a standard. C_1 was varied between 4pF and 400pF and the breakdown thresholds were determined. Since the charge in the capacitor is given by Q = CV, a relationship between the breakdown threshold and the charge can be found. Both positive and negative polarity pulses were applied.

2.5 Continuous Voltage Breakdown Thresholds

Under continuous voltage conditions, the influence of the dwell-time of the applied voltage on the breakdown threshold was determined using the programmable voltage source. Threshold breakdown voltages for dwell-times ranging between 2 ms and 100 ms were determined for samples of 5 capacitors at each time interval. 2 ms was the low limit of the voltage dwell-time available on the power supply. As noted previously, only negative polarity voltages were applied.

3. RESULTS

3.1 Pulsed Voltage Breakdown Thresholds

The pulsed voltage breakdown thresholds as a function of the magnitude of the discharge capacitor are shown in Figure 5 for positive voltages, and in Figure 6 for negative voltages.

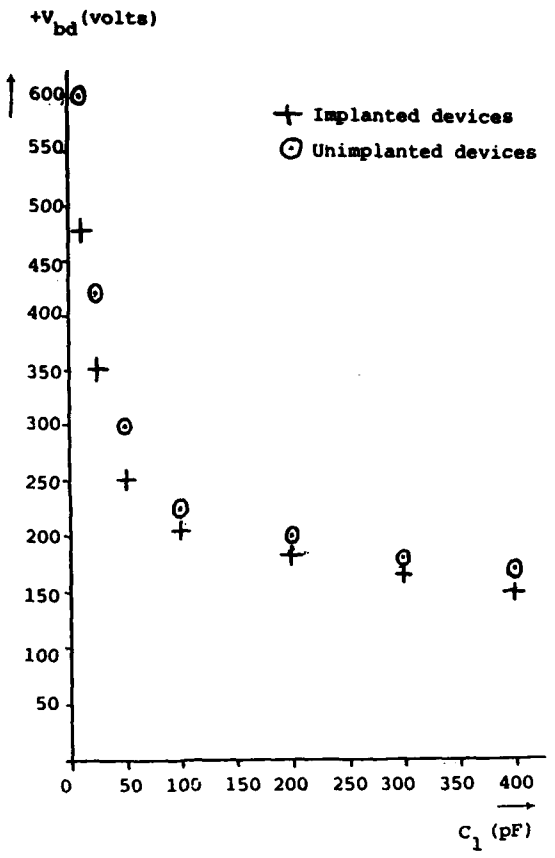


Figure 5 Graph of pulsed V_{bd} vs. discharge capacitor, C_1, for positive voltages

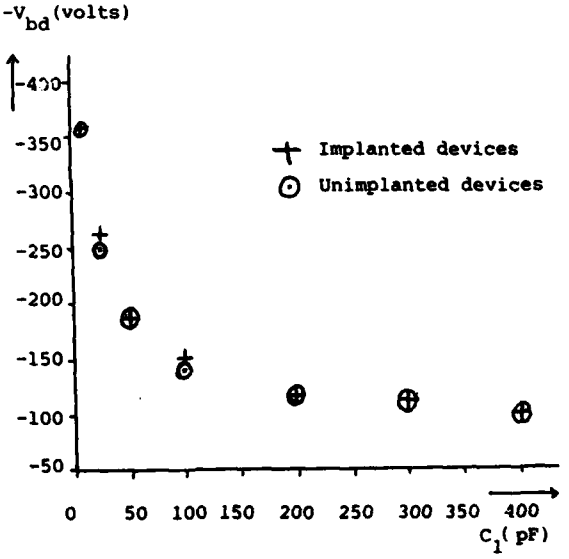


Figure 6 Graph of pulsed V_{bd} vs. discharge capacitor, C_1, for negative voltages.

Each figure indicates the variation of breakdown thresholds for devices which have a preferential boron implant at the Si surface as well as the thresholds for unimplanted devices.

It can be seen from these curves, that the positive voltage breakdown thresholds are much higher than the negative voltage thresholds. They are also different functions of the discharge capacitance. These factors will be discussed in detail in Section 4.

The results proved to be very consistent between devices of identical structures. Errors in the measurements were introduced by the minimum voltage increment, ΔV, available in the static discharge equipment. Typically for positive voltage pulses, ΔV was 20V, indicating an accuracy of ±10V and for negative voltage pulses ΔV was 10V, indicating an accuracy of ±5V.

3.2 Continuous Voltage Breakdown Thresholds

The variation in continuous voltage breakdown thresholds for unimplanted p-Si MOS structures is given in Table 1, as a function of the duration of the applied voltage, Δt. Only negative polarity voltages were used.

Accuracy of the voltage breakdown readings were around ±0.15V limited by the increments possible in determining breakdown for a given Δt. However, great consistency of results were obtained which allows much confidence to be placed in these readings.

TABLE 1

V_{bd} as a Function of Voltage Dwell-Time (Δt) for Continuous Voltage Stress Conditions, for Unimplanted p-Si MOS Capacitors

V_{bd} (Volts)	-36.5	-37.0	-37.75	-38.0	-38.25	
Δt (ms)	100	20	10	5	2	2

4. DISCUSSION

4.1 Dielectric Breakdown in MOS Capacitors

4.1.1 Negative Voltage Conditions

When a negative voltage is applied to the p-Si MOS structure, the gate electrode becomes the cathode. The n^+ polysilicon gate has a very high density of electrons ($\approx 10^{21} cm^{-3}$) available for injection into the SiO_2. In these conditions, the applied field enhances the injection mechanisms by which electrons enter the SiO_2 conduction band, and then provides the necessary energy to cause the SiO_2 to breakdown.

4.1.2 Positive Voltage Conditions

The application of a positive voltage to the gate of the structure causes the p-Si surface of the capacitor to be the cathode. However, the p-Si surface has to first become inverted before electron injection can take place. The response times of minority carriers to an applied voltage are typically between 10^{-3} and 10^{-2} secs [9] [10], and voltages with rise times faster than this would not allow the inversion layer to form. As a result, the maximum depletion layer width increases until the field in the Si is high enough to cause avalanche breakdown. Minority carriers then flood the Si surface [11]. Electrons reaching the Si-SiO_2 interface can have sufficient energy to be injected into the SiO_2. If the injected carrier density is of the right order of magnitude, destructive breakdown of the oxide will occur.

Typical avalanche breakdown fields are in the region of $400 kVcm^{-1}$ depending on the dopant concentration [12]. Prior to avalanching, the field across the MOS structure is divided between the SiO_2 and the Si depletion layer. After avalanche takes place, the whole field appears across the SiO_2 and will contribute to the breakdown process.

4.2 Analysis of Experimental Results

4.2.1 Positive Polarity Pulsed Voltages

V_{bd} for positive pulses is shown in Figure 5, to be inversely proportional to the value of the discharge capacitance, C_1. The pulsed voltage breakdown threshold in the positive case

is dependent on the avalanche breakdown field in the p-Si, as discussed in Section 4.1. Figure 7 schematically depicts the p-Si MOS capacitor under the influence of a large positive electric field.

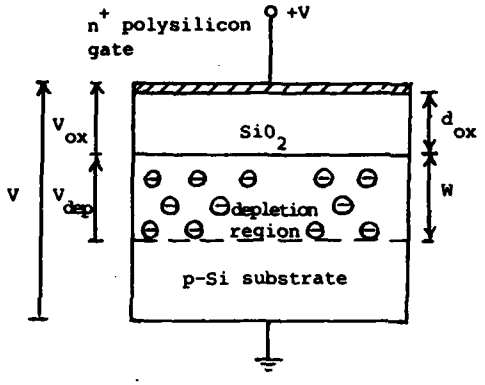


Figure 7 p-Si MOS capacitor under the influence of a large positive voltage

The voltage across the MOS capacitor, V, is made up of 2 voltages V_{ox} and V_{dep}, across the oxide and the depletion layer respectively. The total charge in the capacitor C_2 is given by:-

$$Q_2 = C_2V \quad (1)$$

Where C_2 is the total capacitance of C_{ox} and C_{dep}, given by:-

$$1/C_2 = 1/C_{ox}+1/C_{dep} \quad (2)$$

Since the system is in deep depletion, Q_2 is given by:-

$$Q_2 = qN_AW \text{ Coulombs cm}^{-2} \quad (3)$$

Where N_A is the acceptor concentration per cm^3 in the p-Si and W is the depletion layer width.

It is important at this stage to point out that the capacitor C_1 is an unreplenishable reservoir of charge. Any charge transferred to C_2 will deplete this reservoir and hence reduce the voltage, V, across the system.

The relationship between V and the voltage to which C_1 is initially charged, V_o, neglecting the voltage dropped across R_2, is given by:-

$$V = V_o-(Q_2/C_1) \quad (4)$$

$$\text{Thus } V = V_o/(1+C_2/C_1) \quad (5)$$

$$\text{Now } C_{ox} = \varepsilon_o\varepsilon_{ox}A/d_{ox} \quad (6a)$$

$$\text{And } C_{dep} = \varepsilon_o\varepsilon_{si}A/W \quad (6b)$$

Where it is assumed that the area of the capacitor is large enough that edge effects may be neglected.

$\varepsilon_o\varepsilon_{ox} = 3.45\text{x}10^{-13}$ Fcm^{-1} is the permittivity of SiO_2.

And $\varepsilon_o\varepsilon_{si} = 1.04\text{x}10^{-12}$ Fcm^{-1} is the permittivity of Si.

Also using:-

$$V = V_{dep}+V_{ox} \quad (7)$$

Where:-

$$V_{ox} = Q_2/C_{ox} \quad (8)$$

We get:-

$$V_o = Q_2(A/C_1)+Q_2(A/C_{ox})+(qN_AW^2/\varepsilon_o\varepsilon_{si}) \quad (9)$$

The depletion layer width, W, is given by the solution to Poisson's Equation in the silicon as:-

$$W = (\varepsilon_o\varepsilon_{si}F_{dep}/qN_A) \quad (10)$$

Where F_{dep} is the field across the depletion region. Thus for an acceptor carrier density of 4×10^{15}cm^{-3} and a breakdown field of 350kVcm^{-1}

$W = 5.7 \times 10^{-4}$cm

And for C_1 = 100pF

V_o = 205V

Which is of the order shown in Figure 5.

It is important to note that when V_o is 205V, and with F_{dep} and W as given above,

V_{dep} = 199.5V

Then using Eq (4) and (7)

V_{ox} = 4V

F_{ox} = 1 MV cm^{-1}

Which is well below the oxide breakdown field strength. Hence, it is seen that only after avalanche breakdown has occurred in the silicon, can the SiO_2 breakdown be initiated.

From Eq (9), the dominant term is $(qN_A W^2 / \varepsilon_o \varepsilon_{si})$, and this term determines the minimum V_o for breakdown for a given N_A. As C_1 is increased, V_o approaches this minimum value since the first term in the equation becomes small. This is indicated at the upper end of the curve in Figure 5. For small C_1, V_o begins to rise more rapidly as the term (Q_2/C_1) begins to influence the total voltage.

W is a function of N_A, and the larger the dopant concentration, the smaller W becomes for a given F_{dep}. This, in turn, reduces V_o. Hence, the lower sensitivity of the preferentially-doped implanted MOS capacitors compared with that of the unimplanted capacitors, as shown in Figure 5.

4.2.2 Negative Polarity Pulses

From Figure 6 it is seen that the breakdown thresholds for negative polarity pulses decrease as the value of the discharge capacitor increases. The relationship is found to be given by:-

$$V_{bd} = (1/C_1)^{\frac{1}{2}}.K \text{ volts} \quad (11)$$

Where K is a constant.

K^2 actually has the units of energy and Figure 6 describes a constant energy contour. The total energy dissipated in the MOS capacitor at breakdown is given by:-

$$E = Q_{bd}V_{bd} \text{ Joules} \quad (12)$$

and

$$Q_{bd} = C_1V_{bd} \quad (13)$$

∴ the total energy dissipated at breakdown is given by:-

$$E = C_1V_{bd}^2 \quad (14)$$

TABLE 2

The Effective Voltage across the MOS Capacitor (V), the Charge on the MOS Capacitor (Q_2), as Functions of the Discharge Capacitor (C_1), and the Applied Voltage (V_o).

C_1 (pF)	10	24	50	100	200	300	400
V_o (Volts)	-310	-200	-140	-95	-70	-60	-50
Q_1 ($X10^{-9}$C)	3.1	4.8	7.0	9.5	15.0	18.0	20.0
V(Volts)	-59.6	-72.7	-76.1	-66.9	-57.8	-52.6	-45.0
Q_2 ($X10^{-9}$C)	2.5	3.0	3.2	2.8	2.4	2.2	1.9
E_{bd} ($X10^{-7}$J)	1.5	2.2	2.4	1.9	1.4	1.2	1.0

The capacitance C_2 = 42pf, and $E_{bd} = Q_2V$. In Table 2 the values of V_{bd} are given after adjusting for the charging of C_2. These figures are then used to calculate E_{bd}. From Table 2:-

$E_{bd} = 1.5x10^{-7}$ Joules

Hence, it is seen that a specific energy E_{bd} is required for breakdown to take place. This would be the minimum energy required by the electrons current in the SiO_2 to cause dissociation when energy losses due to collisions and heat loss by homogeneous thermal conduction have been considered (see Section 4.4).

Comparing $C_1V_{bd}^2$ for positive polarity pulses (Section 4.2.1) with $C_1V_{bd}^2$ for negative polarity pulses, the energy required for positive V_{bd} is higher than for negative V_{bd}. Furthermore, in the positive polarity case, the energy is not a constant. However, this is to be expected when considering the conditions in which breakdown takes place in the two cases.

4.2.3 The Variation of V_{bd} with Δt for Continuous Voltage Stress Breakdown

Only injection from the gate electrode was used in this experiment to determine the relationship between Δt and V_{bd}. Larger Δt indicates larger charge densities as given by:-

$$Q = J\Delta t \ Ccm^{-2} \quad (15)$$

And from the energy relationship, where:-

$E_{bd} \propto Q_{bd}V_{bd}$ Joules

which should be a constant for a given MOS structure.

$$\therefore V_{bd} = 1/\Delta t \quad (16)$$

and Δt is inversely proportional to V_{bd} as shown in Table 1. Since J is a time-independent function of the applied field, V_{bd} is directly related to the injected charge Q_{bd}

4.3 Charge Conduction Processes

Three types of charge conduction processes can be identified. They are; (i) tunnelling [13], (ii) Schottky Emission [13] [14], (iii) Poole-Frenkel Emission [15]. In addition, injected charge in the presence of a high electric field can cause avalanche multiplication of electrons as a result of collision ionization in the SiO_2. These added electrons would then gain energy from the field, and hence contribute to the breakdown process.

However, SiO_2 is a wide bandgap insulator, and the potential barrier (ϕ_B) at the electrodes is around 3.2eV. Calculations for the Schottky current, the Poole-Frenkel current and the avalanche current, suggest that these contributions to the overall current density in the SiO_2 are negligible. Solutions for the tunnelling current, J_{tun}, on the other hand, do show that currents of the order required to cause breakdown are possible for the fields concerned.

The tunnelling equation for an image-force lowered barrier potential is given by [16]:-

$$J_{tun} = (q^3/8\pi h)(F^2/\phi_B(1/t^2(y)).\exp\{(-4\sqrt{2m^*}/3\hbar q)(\phi_B^{3/2}/F)v(y)\} \quad (17)$$

Where q is the electronic charge in Coulombs, h is Plank's constant, m* is the effective mass of electrons in SiO_2, and $y = \Delta\phi/\phi_B$

And $\Delta\phi = (q^3F/\varepsilon_o\varepsilon_{ox})^{\frac{1}{2}}$ is the image-force lowering of ϕ_B.

v(y) and t(y) are tabulated functions of y.

For $F = 9.5MV\ cm^{-1}$, (i.e. V = 38.0v), and $\phi_B = 3.15eV$, and taking $m^* = m_o = 0.9\times10^{-30}$kg.

$J_{tun} \simeq 90 \times 10^{-3}\ A\ cm^{-2}$

The associated charge for an area of $49 \times 10^{-5}\ cm^2$ and $\Delta t = 5ms$

$Q_{tun} = J_{tun}\ \Delta t\ C\ cm^{-2}$

Thus $Q_{tun} = 2 \times 10^{-7}\ C$

Q_{tun} is greater than the charge of $\simeq10^{-9}$ Coulombs which brought about breakdown under pulsed conditions (Section 4.2.2). Hence, the tunnelling current can provide sufficient carriers to initiate oxide breakdown.

The energy supplied by these charge carriers is:-

$E = 8\times10^{-6}$ J

which is greater than the energy required for SiO_2 breakdown.

4.4 Energy Dissipation for Dielectric Breakdown

Charge carriers moving under the influence of the applied electric field may gain sufficient energy to result in thermal breakdown by means of Joule heating in the material [2] [16] [17]. It has also been suggested that electrons, when leaving the SiO_2 conduction band, either at the electrode or at a deep trap, will lose sufficient energy to the lattice to create defects and cause damage [18].

The calculated energies obtained in the work presented in this paper are well within the energies required to breakdown SiO_2. However, it has not been possible to calculate a threshold energy for breakdown, due to the unknown factors involved in the calculations.

When considering the time factors involved in breakdown, it is clear that the heat dissipated by homogeneous thermal conduction in the material must be considered. The heat dissipated is given by:-

$$E_{DIS} = K(dT/dx)\ Jcm^{-2}s^{-1} \quad (18)$$

$$\therefore E_{DIS} = K\ (dT/dx).\Delta t \quad Jcm^{-2} \quad (19)$$

Where $K = 15\times10^{-3}\ Jcm^{-1}s^{-1}K^{-1}$ is the thermal conductivity of the SiO_2. dT/dx is the temperature gradient and Δt is the duration of

the pulse. Therefore, it is seen that as Δt decreases, E_{DIS} is reduced. As a result of this, the energy required to cause breakdown is reduced. This has been shown for the pulsed and continuous voltage conditions discussed in Section 4.2 and 4.3.

5. CONCLUSIONS

Dielectric breakdown of p-Si MOS capacitor structures has been discussed. The experimental results have shown that the threshold voltages for breakdown are a function of the charge supplied to the MOS structure.

Pulsed voltage breakdown under positive and negative polarity conditions show a marked difference between V_{bd}. It has been shown analytically that since a positive voltage pulse results in deep-depletion of the p-Si surface, oxide breakdown occurs only after avalanche breakdown has taken place in the silicon. Hence, over 95% of the applied voltage in the pre-breakdown stage falls across the p-Si, with only 5% across the oxide. Negative voltage pulses on the other hand, cause electron injection from the n^+-polysilicon gate; a process which is almost instantaneous. The energy required to cause breakdown is found to be constant.

The discharge capacitor in the pulse generator circuit is an unreplenishable reservoir of charge. The variation of the effective voltage seen by the MOS devices has been calculated as a function of the values of the discharge capacitor and the MOS capacitor. Hence, it is shown that with negative voltage pulses, the effective voltage across the MOS system is less than the applied voltage.

Under continuous voltage stress conditions, a dependence of V_{bd} on the duration of the applied voltage has been observed. It is suggested that as Δt increases, the voltage necessary to generate the energy required for breadown will decrease proportionally.

ACKNOWLEDGEMENTS

We wish to thank Plessey Research (Caswell) Ltd., for the wafers used in these experiments. Our thanks also to Prof. A.K. Jonscher for discussions on this subject.

This paper is published with the permission of the MOD CVD (PE) and we are grateful to them for the sponsorship.

REFERENCES

[1] Harari E., Dielectric Breakdown in Electrically Stressed Thin Films of Thermal SiO_2. - J.Appl.Phys., 49, pp.2478-2489, 1978.

[2] Jonscher A.K., Lacoste R., On a Cumulative Model of Dielectric Breakdown in Solids. - IEEE Trans.on Elec.Insul., 19, pp. 567-577, 1984.

[3] Wolters D.R., Van Der Schoot J.J., Dielectric Breakdown in MOS Devices. Part II: Conditions for the Intrinsic Breakdown. - Philips J.Res., 40, pp. 147-163, 1985.

[4] Olivo P., Ricco B., Sangiorgi E., Electron Trapping/Detrapping within Thin SiO_2 Films in the High Tunnelling Regime. - J.Appl. Phys., 54, pp. 5267-5276, 1983.

[5] Osburn C.M., Weitzman E.J., Electrical Conduction and Dielectric Breakdown, in Silicon Dioxide Films on Silicon. - J. Electrochem.Soc., 119, pp. 603-609, 1972.

[6] Klein N., Mechanisms of Electrical Breakdown in Thin Insulators - an open subject. - Thin Solid Films, 100, pp 335-340, 1983.

[7] Amerasekera E.A., Campbell D.S., Oxide Breakdown in MOS Structures under ESD and Continuous Voltage Stress Conditions. - Proc.Eur.Rel.Conf. (RELCON), pp.325-330, Copenhagen, 1986.

[8] Amerasekera E.A., Campbell D.S., Pulse and Continuous Voltage Breakdown in MOS Capacitor Structures. - Proc.8th Ann. EOS/ESD Symp., Las Vegas, 1986.

[9] Nicollian E.H., Brews J.R., MOS Physics and Technology, John Wiley and Sons (New York), 1982.

[10] Rusu A., Bulucea C., Deep-Depletion Breakdown Voltage of Silicon-Dioxide/ Silicon MOS Capacitors. - IEEE Trans.Elec. Dev., 26, pp 201-205, 1979.

[11] Nicollian E.H., Goetzberger A., Berglund C.N., Avalanche Injection Currents and Charging Phenomena in Thermal SiO_2. - App.Phys.Lett., 15, pp. 174-177, 1969.

[12] Goetzberger A., Nicollian E.H., Transient Voltage Breakdown due to Avalanche in MIS Capacitors. - App.Phys.Lett. 12, pp. 444-446, 1966.

[13] Hill R.M., Single Carrier Transport in Thin Dielectric Films. - Thin Solid Films, 1, pp.39-68, 1967.

[14] Emtage P.R., O'Dwyer J.J., Richardson-Schottky Effect in Insulators. - Phys.Rev. Lett., 16, pp. 356-358, 1966.

[15] Jonscher A.K., Electronic Properties of Amorphous Dielectric Films. - Thin Solid Films, 1, pp. 213-234, 1967.

[16] O'Dwyer J.J., The Theory of Electrical Conduction and Breakdown in Solid Dielectrics. - Clarendon Press (Oxford), 1973.

[17] Stratton R., The Theory of Dielectric Breakdown in Solids. - Progress in Dielectrics, Vol. 3, pp. 235-292, 1961.

[18] Wolters D.R. Van Der Schoot J.J., Dielectric Breakdown in MOS Devices - Part III: The Damage Leading to Breakdown. - Philips J.Res., 40, pp. 164-192, 1985.

TRENDS IN NON-VOLATILE MEMORY DEVICES AND TECHNOLOGIES

H.E. Maes, J. Witters and G. Groeseneken

IMEC, vzw
Kapeldreef 75
B-3030 Leuven, Belgium

In this paper the different technologies and devices used in Non-Volatile Electrically Erasable PROM memories are reviewed, discussed and compared with emphasis on their operation mechanisms, their advantages and limitations, the problems related to further scaling and the reliability issues.

1. INTRODUCTION

The trend in Non-Volatile Memories toward higher bit densities, smaller cell sizes (vertically as well as horizontally), faster access times, lower power consumption, 5 Volt only operation and other user-friendly features is imposed by the requirements of microcomputer systems. Adding reprogrammability allows in-system programmable memories and means increased system flexibility which opens a broad new field of applications such as intelligent controllers, self-adaptive, reconfiguring and remotely adjustable systems, programmable/adaptable logic, artificial intelligence a.s.o. [1].

Whereas 5 years ago at the time of the introduction of 16 kbit Electrically Erasable PROM's (EEPROM) parts, many analysts optimistically projected that by 1990 EEPROM's would have displaced UV-Erasable PROM's (EPROM) as the standard program storage medium in microprocessor based systems, the EEPROM market has not taken off as expected. Figure 1 compares projections made in 1984 and 1986 for the worldwide

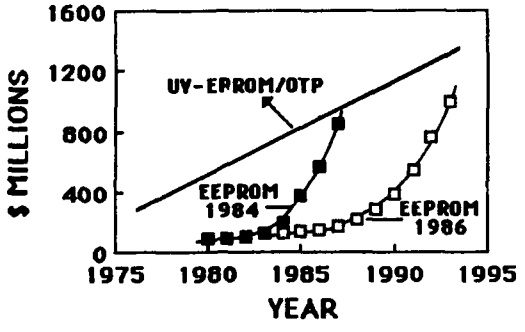


Figure 1. Comparison of estimates in 1984 and 1986 of the global sales of EPROM and EEPROM

EEPROM and EPROM market. Whereas the UV-EPROM market had seen a steady growth, the predicted rise for EEPROM's in 1985 has not been realized. New estimates made in 1986 however predict a definite break-through by 1990. The reasons why the EEPROM market has not grown as predicted have mainly to do with the high cost of the EEPROM's as compared to EPROM's of the same memory density and the lack of real large scale applications for EEPROM. The higher cost is greatly due to the larger cell size of EEPROM's (typically a factor of 4). The traditional EEPROM cell needs at least 2 transistors, a floating gate structure for programming and a select device. The result is that as far as density is concerned conventional (full feature) EEPROM's are at least one generation behind single transistor EPROM's, as is shown in figure 2.

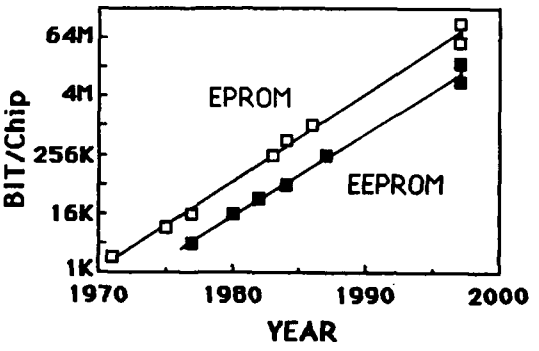


Figure 2. Present and expected level of integration of EPROM and EEPROM parts

However, in their search for non-volatile devices that would still be electrically erasable but as cost-effective as EPROM, manufacturers have come up recently with single transistor EEPROM solutions which combine concepts of EPROM and conventional EEPROM and are called *Flash-EEPROM* [2,3] because all the memory contents of the device are erased at the same time. Whereas the cell size is now comparable to that of UV-EPROM, the device still retains a lot of

EEPROM features although endurance remains unproven and is probably limited to under 1000 cycles. The cost of Flash-EEPROM will probably also become comparable to that of UV-EPROM because a plastic package can now be used instead of the expensive package with a window. The 1986 prediction shown in figure 1 therefore seems more realistic considering that the Flash parts will probably make out the majority of the EEPROM products. At the same time the memory density for EPROM and FLASH-EEPROM will probably become equal.

The *electrically programmable* non-volatile memories can be divided as follows :

(i) UV-erasable EPROM and the One Time Programmable (OTP) device, the latter being contained in a plastic package and therefore not reprogrammable.

(ii) EEPROM parts which can be subdivided into three classes, the all-feature EEPROM, the high speed EEPROM and the Flash-EEPROM. The fully featured part offers all kinds of user-friendly features [1] and is the most sophisticated from the point of view of circuit design. The high speed EEPROM's often offer less features (also smaller densities) but try to compete with bipolar products in speed with access times in the range of 30-50 nsec but of course with lower power consumption.

(iii) Non-Volatile RAM (NOVRAM) which combines the non-volatility of the EEPROM with the ease of use and fast programming of static RAM. The content in the RAM section is transferred to the non-volatile EEPROM portion at power down or power failure.

The state-of-the-art technologies for non-volatile semiconductor memories are the SNOS (silicon-nitride oxide-silicon) and the floating gate type memories. In this paper these two technologies and the different cell structures, cell concepts and operation mechanisms are reviewed and compared. The emphasis in this paper is on the *device aspects of non-volatile memories.* Although as discussed before, EPROM remains the non-volatile memory market leader, its development has been rather straightforward up to the 1 Mbit level in line with the applied scaling practices, with the most advanced 1 Mbit EPROM today using 1-1.2 µm CMOS technology [4]. Therefore this paper will mainly treat the EEPROM technologies and devices which are the most sophisticated in their physical principle of operation. For the EPROM devices only some new approaches necessary for higher density parts will be discussed.

First in section 2 the different *EEPROM technologies and devices* will be discussed separately with emphasis on their operation mechanisms, their advantages and limitations and the problems related to further scaling. A comparison will be made between the different technologies and approaches and new alternative cell concepts will be presented. In section 3 the *reliability issues* related to the different technologies and devices will be treated and finally some perspectives will be given.

2. EEPROM TECHNOLOGIES AND DEVICES

2.1 The SNOS device

MNOS memories were invented 20 years ago [5] and are in fact the first truly electrically alterable non-volatile semiconductor (EAROM) devices. By 1975 p-channel MNOS EAROM with densities of up to 8 kbit were available which used a 1 transistor per bit configuration based on the so-called trigate transistor cell concept [6]. These memories suffered from low speed, limited density, inherent read disturbance and the need for 2 or 3 voltage supplies. An important break-through was achieved for MNOS in 1980 with the development of the n-channel SNOS process [7] which resulted in the first 16 kbit SNOS-EEPROM [8]. The reliability of the SNOS technology was mainly based on the use of Low Pressure CVD silicon nitride and a *pre-metallization high temperature hydrogen anneal [9-12].*

A cross section of the SNOS cell which is still used in todays 64 and 256 kbit SNOS EEPROM parts is shown in figure 3. A two transistor per bit configuration is used where the MOS transistor acts as the select device whose implementation has allowed to completely eliminate the problem of read disturbance [8]. The SNOS transistor consists of a silicon nitride layer (20-40 nm) on top of an ultra thin oxide (UTO, 2 nm) on silicon. Because the SNOS transistor is in fact a two polarity device, necessitating the application of memory bulk voltages, an isolation of the memory bulk from the peripheral circuitry bulk is required. The most common approach is to use separate p-wells for the peripheral MOS circuits and the memory array. Providing separate wells within the memory array will then allow full byte function.

The non-volatile function of the SNOS device is based on the charge storage in discrete traps in the nitride layer. Net positive or negative charge can be stored in almost equal amounts in LPCVD nitrides. The programming of the memory cell is as follows: during the write operation a high voltage is applied to the gate while the well is grounded. Electrons will tunnel from the silicon conduction band into the silicon nitride

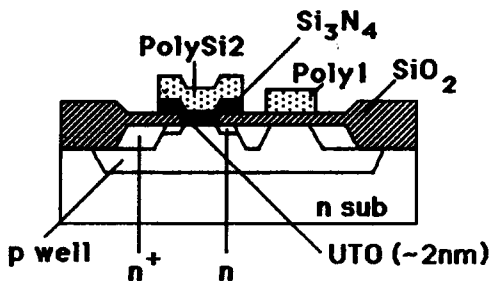


Figure 3. Cross section of SNOS cell used in 64 kbit EEPROM

conduction band by modified Fowler-Nordheim tunneling and will be trapped in the nitride traps resulting in a positive threshold voltage shift. Erasing is achieved by grounding the gate and applying a high voltage to the well. This will induce direct tunneling of holes from the silicon valence band into the nitride traps [13-14] resulting in a negative threshold voltage. During the off-state, the gate is grounded and the select transistor is required for proper operation within the array. Reading of the cell is done by addressing the cell through the select transistor and by sensing the state of the SNOS transistor. Although the gate is grounded, the charge content within the nitride will be modified in time due to charge leak through the UTO oxide (see section 3).

It was shown by Hagiwara et al [14] that the integrity of nitride layers can be guaranteed to thicknesses below 20 nm. Scaling of SNOS devices which must be carried out in parallel to the scaling down of the peripheral MOS transistors is straightforward up to memory densities of about 256 kbit. Table I shows the scaling scheme of Yatsuda [13-15] which is based

TABLE I. Scaling scheme of SNOS

Item	16 k (k=1)	64 k (k=1.5)	256 k (k=2.5)	SF
TECHNOLOGY (μm)	3	2	1.2	k^{-1}
CELL AREA (μm^2)	400	180	60	k^{-2}
MOS : t_{ox} (nm)	75	50	30	k^{-1}
SNOS : t_n (nm)	50	32	20	$k^{-1.1}$
: t_{ox} (nm)	2.1	2.0	1.9	$k^{-0.1}$
Program. Voltage (V)	25	16	10	$k^{-1.1}$

on a reduction of the dimensions almost in proportion to the program voltage except for the UTO. In order to conserve a constant programming time the UTO thickness has to be reduced slightly. A 256 kbit version with a 1.2 μm technology has a 20 nm thick nitride layer and a programming voltage of only 10 V. Hole injection from the gate can however seriously limit the memory window [13,16-18], a problem which becomes more severe for thinner nitride layers. An efficient way of blocking this injection is to provide a thin oxide (2-3 nm) on top of the nitride yielding the so-called SONOS device [19] (figure 4b vs 4a). This oxide can be formed by steam oxidation of the nitride at the expense of the nitride thickness, or by deposition. The blocking efficiency of this top layer has been proven for both types of oxide [13,19,20]. However, when thinning the

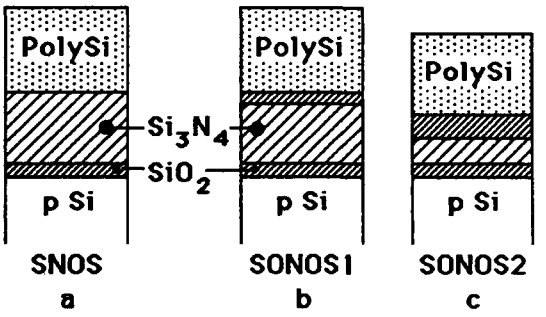


Figure 4. Scaling concepts for SONOS devices

nitride below 20 nm another problem arises. It is well known that the trapping length in nitrides is larger for holes than for electrons, i.e. typically 15-20 nm for holes [21,22] and 5-10 nm for electrons [22,23]. When reducing the nitride thickness holes will therefore be trapped close to the gate and will in fact be mostly lost through the gate electrode, even in the presence of the thin top oxide. This will result in a significant reduction of the threshold voltage in the erased state.

Therefore further scaling of the SNOS device for higher density memories and lower programming voltages will require a new device concept. This has been proposed by Suzuki et al [24] and is now in development [25,26]. The new SONOS concept is shown in figure 4c and consists of a UTO of the same thickness as before, a thin nitride layer(< 10 nm) and a thicker top oxide (> 3 nm). The aim of the latter is not only to inhibit gate injection but also to block the charges injected from the silicon at the top oxide-nitride interface resulting in a higher trapping efficiency and thus reducing the problem related to nitride layer reduction. In this way the total thickness of the insulator structure can be reduced and consequently also the

programming voltage. The new device shows additional advantages : larger oxygen related electron trap densities at the nitride-top oxide interface due to the oxidation of the nitride [27] and which will result in a larger memory window in spite of the decreased nitride thickness. For a constant top oxide layer thickness this will eventually make the threshold of the written state independent of the nitride thickness [26]. Next if pinholes would be present in the thinner nitride layer, they would be filled with oxide afterwards during the oxidation of the nitride. Finally the retention and degradation behaviour are improved as will be discussed in section 3. Low voltage operation down to 5 V has already been demonstrated [28] for a nitride of 3 nm and a blocking oxide of 5.5 nm. It seems that although optimization of the process and structure is still required, the application of this SONOS cell concept will allow the realization of memories with densities in the Mbit range.

Finally SNOS memories have two features which are worth mentioning since they have made this technology the first choice in military and space applications requiring non-volatility. The first one is their inherent radiation hardness and the second the adjustable nature of their memory properties. SNOS devices are inherently hard essentially because unlike silicon dioxide, the mobilities of electrons and holes are not much different in nitrides , so that when exposed to ionizing radiation both generated carriers can be swept rapidly out of the insulator resulting in a negligible amount of trapped charge [29]. Acceptable shifts for total radiation dose up to Megarad (Si) at 77 °K have been obtained for SNOS structures [30]. Reduction of the programming voltages and the abscence of thick oxide parts in the newer SNOS devices has improved their radiation hardness. Anyhow, special techniques are required and have been developed to harden the SNOS/MOS or SNOS/CMOS combination [31,32]. For the new scaled down SONOS devices discussed above, the inherent radiation hardness is however a matter of concern in view of the increasing dominance of the SiO_2 parts in this device . For the first generation SONOS devices (figure 4b) however radiation hardness is still preserved [33].

The second feature is that SNOS devices can be adjusted to the envisaged application : very slow programming (1-100 msec) for long non-volatile retention (years, EEPROM function) or fast programming (1-10 μsec) for limited non-volatile retention (hours-days, NVRAM function) [33].

2.2 Floating gate memory devices.

The first floating gate memory was proposed in 1967 [34]. The first working devices were developed in 1971 [35] ; the programming was based on the injection of high energetic electrons from an avalanche plasma in the drain region. However this programming process is inefficient and erasure was only possible by UV or X-ray irradiation. These first memory devices were all p-channel devices ; in n-channel devices, avalanching the drain only yields hole injection, which is even far less efficient. This has found application [36], but this approach has been abandoned soon in favour of other, less power consuming programming mechanisms.

Nowadays, three different mechanisms are used to change the amount of charge on the floating gate : *Fowler-Nordheim tunneling* (FN) through thin oxides (<12 nm) [37], *enhanced Fowler-Nordheim tunneling through polyoxides* [38,39] and *channel hot electrons* (CHE) [40,41]. Fowler-Nordheim tunneling requires injection fields of the order of 10 MV/cm to narrow the Si-SiO_2 barrier so that electrons can tunnel from the Si conduction band into the SiO_2. In thin oxides the injection field is equal to the average field in the SiO_2 whereas in polyoxides the field at the injecting interface is much larger than the average oxide field due to field enhancement at asperities on the polysilicon surface. Average oxide fields of the order of 2 MV/cm are sufficient to yield injection fields of the order of 10 MV/cm. However, the injection will be enhanced only for one tunnel direction. Channel hot electrons are created in the high electric field near the drain junction. If the oxide field favours injection, electrons can be transported from the substrate to the floating gate. CHE injection can only bring electrons onto the floating gate ; by tunneling, electrons can also be removed from the floating gate. Figure 5 shows the six possible combinations that could be devised to obtain non-volatile floating gate memories and indicates how these devices are programmed. Table II summarizes all their features.

Thin tunnel oxides are difficult to achieve without defects. The growth of textured polyoxides has to be carefully controlled in order to obtain the desired interface features (shape, size of the asperities) which determine the injection current characteristics. Wearout features are much dependent on the achieved polysilicon-SiO_2 interface. CHE injection finally can be optimized by adjusting the substrate doping. However, degradation of the transistor characteristics is unavoidable with this programming mechanism [42]. Memory cells that use a

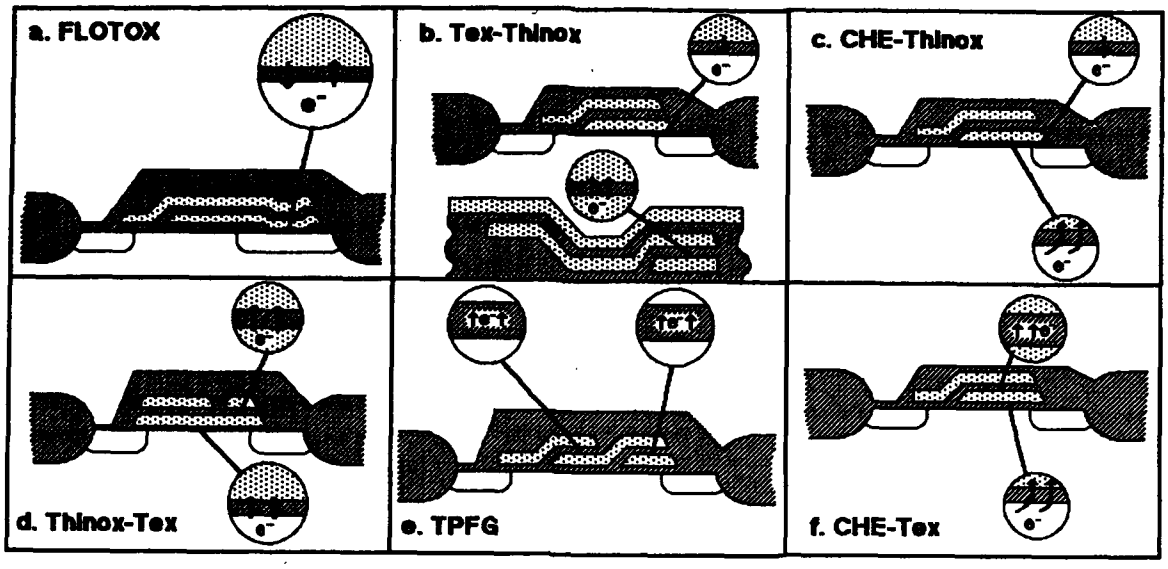


Figure 5. The six possible non-volatile floating gate memory structures and the way they are being programmed.

Table II Features of the six memory cells of figure 5.

a FLOTOX EEPROM/ASIC/Logic + / -	*b* Tex-Thinox EEPROM/F-EEPROM + / -	*c* CHE-Thinox EEPROM/F-EEPROM/ASIC + / -
• compatability • low development entry cost • possible with 1 poly layer • large cell • difficult scaling • thinox defect density	• small cell • easily scalable • low progr. power • complex cell • 2 progr. mechanisms -thinox (defects) -polyoxide (wearout)	• small cell • relativly simple cell • easily scalable • high progr. power • 2 progr. mechanisms -thinox (defects) -CHE (degradation) • ?? endurance ??
d Thinox-Tex EEPROM/F-EEPROM + / -	*e* TPFG EEPROM/ASIC + / -	*f* CHE-Tex EEPROM/EPROM/ASIC + / -
• low progr. power • large cell • scaling difficulties • 2 progr. mechanisms -thinox (defects) -polyoxide (wearout)	• thick oxide • small cell • low progr. power • direct write • easily scalable • complex cell • higher prog. voltage -trap-up -window variation • critical tunnel oxide	• small cell • thick oxide • easily scalable • high progr. power • 2 progr. mechanisms -polyoxide (wearout) -CHE (degradation)

different mechanism for erasing and for writing of the cell, are technologically more difficult to optimize.

Because for memory cells that use thin oxides, the injection field equals the average oxide field, these cells need a large coupling between the floating gate and externally controlled terminals of the device. The high gate capacitance of the sense transistor can be used for this purpose during the programming operation that lowers the threshold voltage. If only this operation uses thin oxide FN tunneling, a small cell size can be achieved [43,44]. But if both programming operations use FN tunneling, large coupling capacitance areas are needed [45]. Polyoxides are much thicker and conduct currents at lower average oxide fields, so that no large coupling capacitance areas are needed [46].

Scaling is often difficult because of the complex cell layouts used. To allow the programming voltages to decrease, thin tunnel oxides should become even thinner. Thicknesses of 6 nm however are the limit for good retention behaviour. But these oxides are hard to grow with low defect densities. Yield considerations now limit the used oxides at 8 to 10 nm [47]. Thinner tunnel oxides imply higher capacitances and thus larger coupling areas. Real small floating gate tunnel oxide (FLOTOX)

cells are therefore hard to achieve. The need for thin polyoxides is not that stringent : the injection current through these oxides is determined primarily by the shape and size of the asperities and not by the oxide thickness [48,49]. The smaller number of injection points of a scaled textured poly floating gate memory cell can however aggravate the intrinsic wearout due to trap-up. Scaling structures using CHE can give new opportunities to this kind of memories. At small channel lengths, a drain voltage of 5 V can be sufficient to generate hot electrons [44]. Only the gate will then need a higher voltage for programming. Because no large currents have to be supplied at this voltage, the circuit could be externally 5 V-only. In the latter , the high programming voltage is generated on chip, by a capacitive voltage multiplier [50]. This *charge-pump* can generate high voltages but because of its very large output resistance, no current at high voltage can be allowed. Non-volatile memories with CHE programming could thus operate from a 5 V supply voltage. But the high programming current remains and therefore it is not likely that conventional EEPROMs (which are reprogrammed frequently) will be constructed in this technology.

Of these six possible floating gate memory cells, two have never been used in commercial products. The *Thinox-Tex* approach seems to have only drawbacks (cfr. Table II) : large cell areas are needed due to the thin oxide Fowler-Nordheim tunneling, used for erasing the cell. This also makes the cell difficult to scale. The two different programming mechanisms enhance the technological problems. The *Tex-Thinox* cell has a rather complex structure : three polysilicon layers are needed, and the two different programming mechanisms have to be optimized separately. This cell has not been used in commercial products either. The *CHE-Tex* device has been used in the past [40,41]. However, the high programming power was incompatible with externally 5 V-only operation. This drawback could disappear, if sufficient scaling is possible : at very short channel lengths (<1μm), a drain voltage of 5 V is sufficient to generate hot electrons for programming [44]. The two different programming mechanisms however cause the used technology to be complex.

The *FLOTOX* cell is used in many commercial products. One of the main reasons for this, is the low development entry cost : the only additional process step in standard double poly processes, is the growth of the thin oxide. It is even possible to realize this type of non-volatile memory in single poly processes. This makes this cell highly suitable for ASIC and logic applications [51,52] : the drawback of the larger cell area is not so important for these applications. For large memories however, further scaling and the use of thin oxides become madatory if memory density exceeds the 16K level [47,53]. The use of new tunnel materials can obviate some problems : oxinitrides or nitrided oxides offer better endurance [54,55] ; oxides grown on highly doped injection regions show higher tunnel current conductance (see section 2.3.).

Textured Poly Floating Gate (TPFG)devices are used in large density memory circuits. The main difficulty in this approach, is the growth of the polyoxides with the desired features. The used structure is rather complex : three polysilicon layers or two layers and an additional burried contact [56], are needed. The accurate alignment of these layers requires precise lithography. This has held back the use of this cell in ASIC or logic applications. The small cell size and the ease of further scaling make this structure however the best choice among the floating gate devices for large density memory applications.

A recently introduced approach, is the *CHE-Thinox* cell (used for flash EEPROM). As this cell relies on thin oxide Fowler-Nordheim tunneling only for lowering the threshold voltage, no large coupling areas are needed. Through scaling one can obtain a 5 V-only product [43,44]. However, the high programming power still remains a drawback. The two different programming mechanisms can introduce technological problems. Endurance could be poor because of the CHE degradation effects ; endurance for the commercial products still remains unproven.

2.3 Variations and possible improvements

As discussed in previous sections, the three basic injection mechanisms for the programming of floating gate EEPROM devices can be and have been applied in various combinations and memory configurations. However, besides these combinations, several variations to the mechanisms themselves or to the cell design have been proposed in order to yield alternative memory cells with an improved performance.

These variations generally have one common purpose, namely to increase the injection currents without increasing the programming voltages. In this section these approaches are reviewed. Whereas some of them are merely new ideas with a possible chance for future application, others have already been implemented in real memory circuits.

The injection of hot electrons is a very inefficient process, due to the fact that a proper biasing condition for hot electron

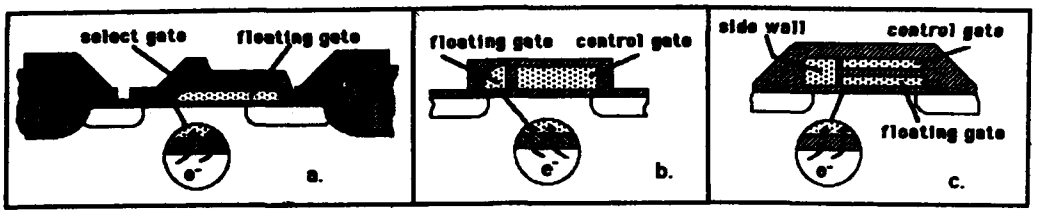


Figure 6. Three alternatives to increase the hot electron injection efficiency. a) PACMOS-cell; b) side-wall floating gate cell; c) source side injection cell.

generation does not go together with a favourable condition for oxide injection. In practice, very high drain and gate voltages are needed, and the injection efficiency is very low. Alternatives should be based on creating a high hot electron generating electric field in the channel and simultaneously a favourable oxide injection field at the site of the hot electron generation. Several solutions that meet these requirements have been proposed.

In one approach this has been achieved by using a dual gate structure, as shown on fig. 6.a [57] . By keeping the first gate, which acts as a select gate, at a voltage close to the threshold voltage, while biasing the second (floating) gate to a high voltage (via capacitive coupling from the drain), a strong potential drop is caused in the center of the channel where neither of both gates are controlling the channel potential. The injection occurs at this site, and simulations have demonstrated that the maximum current path is directed perpendicular to the interface. The injection efficiency could be increased from 10^{-7} for the conventional hot electron injection to 10^{-3} for the PACMOS cell (perpendicularly accelerating channel injection MOS).

Another approach uses a side-wall floating gate structure, as illustrated on fig. 6.b [58]. Instead of controlling the whole channel region, the side-wall floating gate is only controlling a small part of the channel. During programming, the junction at the side-wall floating gate is used as the drain, in order to inject electrons on the floating gate. During reading, the same junction is used as source, and the floating gate is now controlling the flow of carriers from this junction into the channel. This has the additional advantage that the cell has a very good tolerance to unexpected programming during reading, because the hot electrons in that case are generated at the side of the channel opposite to where the floating gate is situated.

A third approach that was proposed to increase the injection efficiency of the hot electron injection mechanism is the source-side injection, shown on fig. 6.c [59]. It uses a side-wall gate and a conventional stacked gate structure. Under the spacer oxide between the side-wall gate and the stacked gates, a weak gate control region is formed. This creates a high channel field, located near the source, where the oxide field is highly favorable for injection. The injection efficiency of this cell is in the order of 10^{-5} to 10^{-6}.

For the tunneling mechanism, several possibilities can be considered to increase the injection efficiency without changing the programming voltage: reducing the oxide thickness, reducing the oxide injection barrier or increasing the coupling factor between control and floating gate without needing too large cell areas.

Reducing the tunnel oxide thickness however puts severe constraints on the device reliability due to direct tunneling leakage problems and the enhanced problem of layer integrity. Reducing the tunnel oxide barrier has been obtained e.g by using Si-rich SiO_2 [60] or nitrides [61]. In fact, also the use of textured poly-oxide, that was already discussed, is one of the possible solutions that has become successful. Another possibility is to grow the thin tunnel oxide on a highly doped n-type Si substrate. In this way, the energy barrier can be reduced from about 3.1 eV down to 1.8 eV, allowing programming voltages of only 12V with tunnel oxides of 14nm [62,63].

Increasing the coupling factor can also be obtained in several ways. One of the possibilities is illustrated on fig. 7 (Shielded substrate injection MOS). The floating gate is shielded from the substrate by the control gate [64]. In this way, the coupling factor from control to floating gate can be increased without needing a large control to floating gate overlap area. At the same time, the control gate between the floating gate and the substrate acts as the serial read transistor, which reduces the cell area even more. In this way, an improved coupling factor and a reduced cell area are combined. This type of cell takes only about a quarter of the area of the larger cells (FLOTOX).

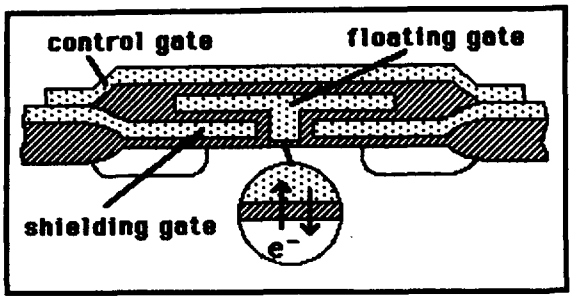


Figure 7. Shielded substrate injection memory cell. The floating gate is shielded from the substrate by the shielding gate, which acts at the same time as control gate.

Another approach to increase the coupling factor is to replace the couple capacitor of the conventional stacked gate structure by one that is formed by a tunnel-oxide MOS capacitor (fig. 8) [51].

Finally, we mention the idea of a single transistor SONOS EEPROM cell, as illustrated on fig. 9 [65]. This is a conventional SONOS cell, but it uses the programming mechanisms of a floating gate cell (hot electrons and hot holes). This leads to several attractive features: injected electrons or holes are only stored in the nitride at the drainside, and consequently no selection transistor is needed. The cell area can be minimal, while the cell has much better retention characteristics than conventional SNOS memory cells, due to the use of a thicker oxide layer.

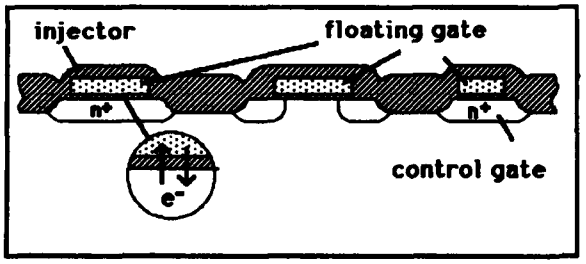


Figure 8. Single poly memory cell. The voltages are coupled to the floating gate via the n^+-diffusion, which acts as control gate.

3. RELIABILITY ASPECTS

3.1. SNOS devices

The two important reliability issues for the SNOS technology are the information retention and the Write/Erase endurance.

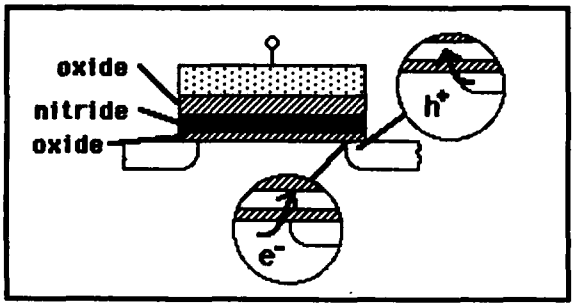


Figure 9. SONOS EEPROM cell. Writing occurs through hot electron injection, erasing by hot hole injection. The need of a selection transistor is avoided.

The threshold voltage of the programmed SNOS device decreases in time, which can be ascribed to loss of net charge in the nitride layer either by backtunneling to the silicon bands [66,67] or by injection from the silicon into the nitride of carriers of the opposite type [68] and/or by redistribution of the charge in the nitride layer [69].

The loss by backtunneling can be reduced by an appropriate hydrogen anneal step [11,14] by which the interface state density is strongly decreased [70]. For conventional SNOS and SONOS devices the threshold voltage decay is logarithmic in time and from extrapolation of the data taken over several decades of time, retention times of well over 10 years can be expected, even at elevated temperatures. Furthermore a slow down of the decay rate is observed for longer times [26,68]. For the scaled down SONOS device it is not clear at present whether the further scaling will really improve the retention behaviour as has been predicted [26]. If in this scaled down SONOS device the major contribution to the threshold voltage shift is made up by the charge stored at the top oxide-nitride interface, backtunneling can indeed be expected to strongly decrease. However in this case injection from the silicon and compensation of the stored charge could become a major issue of concern [68] in view of the high (> 2MV/cm) fields that will be reached in the tunneling oxide for stored charges in the order of 10^{12}-10^{13} cm^{-2}.

Endurance is defined as the ability of the device to perform as specified without degradation when subjected to repeated Write/Erase cycling. In conventional SNOS devices it has been shown that the degradation is primarily caused by hole transport through the tunneling oxide [71,72] by which hole traps are created in the oxide and interface states are generated [70,73], resulting in shifts of the threshold voltage and reduced

retention. It was shown recently that the hole transport towards the silicon is the most damaging [72]. This explains why reduction of hole injection from the gate in SONOS devices has given rise to improved endurance [13]. For the scaled down SONOS structure, hole injection is almost eliminated and consequently even better endurance can be expected. This has been confirmed recently on ultrathin SONOS structures (2 nm SiO_2, 8.5 nm Si_3N_4, 5 nm SiO_2) which showed no noticeable degradation after 10^7 10 V (E/W) pulse cycles [74].

3.2. Floating gate devices.

Under 5 V operation non-volatile memories show a failure rate below other semiconductor memories [47] : non-volatile technologies are designed to handle high voltages. High voltage screening (simply cycling the devices a few times) is sufficient to detect defect gate oxides. Floating gate memories suffer from a lower failure rate due to α-particles than volatile memories, since the charge is stored on a floating gate which is less susceptible to α-particle induced electron-hole pairs [75]. Reliability of non-volatile memories is thus foremostly determined by the non-volatile programming operations.

For floating gate devices there is no intrinsic retention problem [47] : retention is only limited by defect densities. Defects can be activated by the stress the oxide layer undergoes in a high temperature bake or from a large number of programming cycles [76]. A lot of endurance failures are thus retention failures.

Program/erase endurance of floating gate devices is determined by 4 phenomena: tunnel oxide breakdown, gate oxide breakdown, trap-up and degradation of the sense transistor characteristics. Whereas the first two are self explaining, *trap-up* is defined as the trapping of electrons in the oxide during programming operations. This causes the injection currents to decrease. *Degradation of the transistor characteristics* occurs when CHE injection is used for programming [42]. CHE injection is mostly used in EPROM where endurance is restricted to a low number of cycles. For the recently introduced EEPROM devices using CHE injection [43,44], no endurance characteristics are yet available. Fowler-Nordheim injection also causes degradation of the transistor characteristics, but most of the devices make use of a separate tunnel area, leaving the sense transistor unaffected by the programming operations.

The reliability problems are now discussed for the two presently used EEPROM approaches. FLOTOX memories show a broad random-life failure rate distribution with an almost flat failure rate [47]; textured poly floating gate devices (TPFG) have fewer random-life failures but a sharper wearout beyond a number of cycles (see figure 10, [47]). The failure rate due to tunnel oxide breakdown is determined by defects in the oxide. A low percentage of the memory cells contain defective tunnel oxides which become increasingly leaky (the cell fails to retain data) with cycling, until breakdown occurs. This phenomenon is responsible for almost all FLOTOX endurance failures [47,53]. The thicker tunnel oxide of the TPFG devices contains less defects ; half of the pre-wearout failures is caused by the oxide breakdown, but the failure rate is far less than for the FLOTOX type devices [47,53].

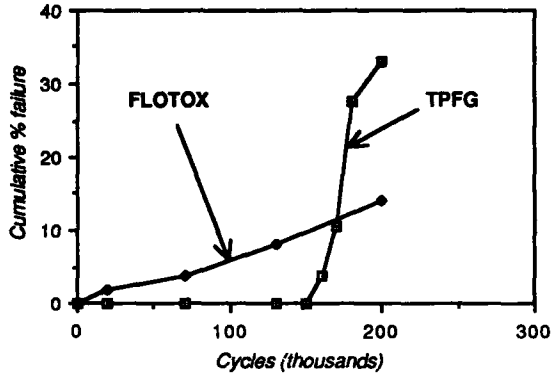


Figure 10. Endurance comparison of FLOTOX and TPFG 16 kBit memories (ref. 47).

Programming puts high voltage stress on both the cell and the logic circuitry, causing defective gate oxides to break down. This accounts for half of the failures before the onset of wearout of TPFG device and some of the early life failures in FLOTOX devices [47,53].

Trap-up is seen to be the dominant endurance limitation in TPFG memories : the thick tunnel oxide combined with the local high tunnel injection current densities and the low average field in the tunnel oxide, are causing a local high trapping rate [48]. Variation in the programming characteristics, even between adjacent bits, must be compensated by a higher programming voltage, causing even more trap-up. This explains the high failure rate after some number of programming cycles. The trap-up reduces however the chance for oxide breakdown, by screening the spots with high conductivity. FLOTOX devices (thin tunnel oxide, high field, uniform injection) suffer less from trap-up.

In scaling FLOTOX devices, the thin tunnel oxide area is reduced, and so are the defect related endurance problems ; on the other hand, thinning the tunnel oxide will again enlarge the problems [47,53]. New materials like oxinitride or nitrided oxides can offer improved integrity of thin tunnel films [54,55]. Scaled TPFG devices will show a larger variation of program/erase voltages due to the smaller number of injecting points per cell. Therefore the programming voltage should be increased, aggravating trap-up problems. Thinning the oxides however will lower the trap-up rate [49].

Due to the defect-related nature of the failures of FLOTOX devices, they will be best suited for low-density applications (small memories, ASIC) [77]. TPFG devices suffer from an intrinsic wearout failure mechanism that, once taken care of, has only a small dependence on memory size (due to the dependence of the worst bit window on the number of bits/chip)[47]. High density memories with the same reliability features can be made. Screening via program/erase cycling is useful for reducing early life failures, but the fairly flat random-life failure rate of the FLOTOX devices prevents this screening from being fully effective [47].

4. COMPARISON OF THE EEPROM APPROACHES AND TECHNOLOGIES

As has become clear from the previous sections, the different technologies and memory approaches have merits and drawbacks. The progress made in the physical understanding of the different programming mechanisms, the mastering of the technology and the capability of adapting the technology and the device to a specific application is such that any technology or approach could be used or engineered for almost any application, provided sufficient effort is put into the development. However other criteria such as the development entry cost, being the amount of extra effort required to bring up a new technology [77], compatibility with standard technology, background in the technology, environmental requirements, a.s.o. are factors that will greatly influence and determine the selection of a particular technology or approach.

In Table III 4 technologies or approaches (SNOS, Textured poly (TP), FLOTOX and Flash-EEPROM using CHE and Fowler-Nordheim tunneling) are compared against a number of criteria.

Table III : Comparison of NV-approaches

Criteria	SNOS	TP	FLOTOX	Flash
• Scaled cell size	+	+	-	++
• Voltage scaling	++	+	-	+
• Complex.of technology	H	H	L	L
• Compatib. with stand. T.	o	+	++	++
• Complexity of cell	L	H	M	M
• Retention	+	++	++	++
• Endurance	+	+	+	o
• Radiation hardness	++	-	-	-
• Entry cost	H	H	L	L

++ : very good ; + : good ; o : medium ; - : poor
H : high ; M : medium ; L : low ;

5. PERSPECTIVES AND CONCLUSIONS

It can be stated that the non-volatile memory technology is an ever evolving field with basically two types of products, the fairly mature EPROM parts and the EEPROM family which is about to reach a maturity level which undoubtedly will be at the basis of a range of new applications. The latter were already foreseen years ago but have not seen realization for different reasons discussed in this paper. The EEPROM's will evolve in two categories, the high density parts for application as data memory and low density (and probably also high speed) units for application in Electrically reprogrammable logic devices (EPLD) and different forms of ASIC's. In the former the fully featured devices can be distinguished from the Flash-EEPROM's which should rather be viewed as flexible EPROM's. From the discussion in this paper it is clear that for each category a best technology or approach will emerge. High density parts, fully featured EEPROM's will most probably only be realizeable with Textured Poly and the SNOS approaches. For both technologies 256 kbit parts are available or will be shortly, and reaching higher densities seems straightforward. SNOS parts will of course be mainly used in military and other applications requiring high radiation tolerance. FLOTOX technology is favoured in low density parts for EPLD and ASIC in view of its compatibility with the used technology. The latter argument and also the higher cost of the textured polysilicon process will probably limit its use in these applications. SNOS is being used in some of these applications [78] but its use will again remain limited. For Flash-EEPROM the different approaches still seem to be applied and although the requirements for the thin tunneling oxide used in the FLOTOX approach are maybe not as stringent as for the fully featured memories, again this might become the limitation for this technology in the long run. At

present, 128 kbit Flash parts have been announced.

The expected market place for the different NV memories in 1987 and 1997 is shown in figure 11 whereas the expected memory size is given in Table IV (see also figure 1)

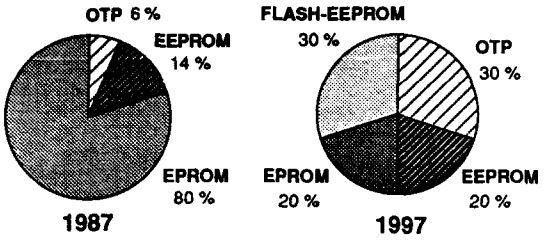


Figure 11. Present Market share of the different NV memories in 1987 and expected share for 1997

TABLE IV : Progress in NV memory density

part	1987	1997
EPROM + OTP	1 Mbit	64 Mbit
Flash EEPROM	512 kbit	32-64 Mbit
All-featured EEPROM	256 kbit	8-16 Mbit

The prognosis is clearly that the Flash-EEPROM will become a major market, mostly at the expense of EPROM (which with OTP however remains the Non Volatile market leader) but also partly at the expense of the full function EEPROM. It is to be expected that system applications of non volatile memories will continue to grow as the density and the reliability of these memories will increase and that the application of EEPROM technology to logic will create a new market of Programmable/Adaptable logic.

ACKNOWLEDGEMENTS

The authors would like to thank R. Pashley from INTEL, F. Masuoka from TOSHIBA and P. Suciu from AMD for their valuable information. J. Witters is a research assistant of the IWONL and G. Groeseneken is a senior research assistant of the Belgian Fund for Scientific Research. Part of this work is also sponsored by the European Space Agency.

REFERENCES

[1] Maes, H.E., Digest Tech.Papers,ESSCIRC83(1983)1
[2] Masuoka, F., Asano, M., Iwahashi, H., Komuro, T.and Tanaka, S., Tech. Digest IEDM 1984 (1984) 464
[3] Masuoka, F., Asano, M., Iwahashi, H., Komuro, T. and Tanaka, S., Digest Tech. papers ISSCC85 (1985)168
[4] Saito, S., Tanaka, S., Atsumi, S., Yoshikawa, K.,Sato, M., Mahita, K., Mori, S., Nozawa, H. and Iizuka, T. , Digest Tech. Papers ISSCC85 (1985) 176
[5] Wegener, H.A.R., Pao, H.C., O'Connell, M.R. and Olehiak R.E., 1967 IEDM Tech. Digest (1967)
[6] Cricchi, J.R., Blaha, F.C. and Fitzpatrick, M.D., 1973 IEDM Tech. Digest (1973) 126
[7] Yatsuda, Y., Hagiwara, T., Kondo, R., Minami, S. and Itoh, Y., Proc. 10th Conf. Solid State Devices (1979),11
[8] Hagiwara, T., Yatsuda, Y., Kondo, R., Minami, S., Aoto, T. and Itoh, Y., IEEE J. Solid State Circuits SC-15 (1980) 346
[9] Schols, G., Maes, H.E, Declerck, G. and Van Overstraeten R., Revue de Phys. Appl. 13 (1978) 825
[10] Yatsuda, Y., Minami, S., Kondo, R., Hagiwara, T. and Itoh, Y., Jap. J. of Appl. Phys., 19, S. 19-1 (1980) 219
[11] Maes, H.E. and Heyns, G.L., J. Appl. Phys., 5(1980) 2706
[12] Schols, G. and Maes, H.E., Proc. ECS Vol 83-8 (1983) 94
[13] Yatsuda, Y., Hagiwara, T., Minami, S., Kondo, R., Uchida, K. and Uchiumi, K., *Jap. J. of Appl. Phys.*, 21, S. 21-1 (1982) 85
[14] Hagiwara, T., Yatsuda, Y., Minami, S., Naketani, S., Uchida, K. and Yasui, T., 6th NVSM, Vail, Aug 1983
[15] Yatsuda, Y., Minami, S., Hagiwara, T., Toyabe, T., Asai, S. and Uchida, K., 1982 IEDM Tech. Digest (1982) 733
[16] Schroder, D.K. and White, M.H., IEEE Trans. Electron Devices ED-24 (1977) 584
[17] Maes, H.E. and Heyns, G.L., 4th NVSM, Vail, Aug 1980
[18] Maes, H.E. and Heyns, G.L., *Insulating Films on Semiconductor 83*, Eds. Verweij J. and Wolters D. (1983) 215
[19] Chen, R.C., IEEE Trans. Electron Devices ED-24 (1977) 584
[20] Remmerie, J., Maes, H.E., Witters, J. and Beullens, W., Ext. Abstr. Fall Meeting ECS 1986, Abstract 374
[21] Maes, H.E. and Van Overstraeten, R., Appl. Phys.Lett. 27 (1975) 282
[22] Hampton, F.L. and Cricchi, J.R., 1979 IEDM Tech. Digest (1979) 374
[23] Yun, B.H., Appl. Phys. Lett. 27 (1975) 256
[24] Suzuki, E., Hiraishi, H., Ishi, K. and Hayashi, Y., IEEE Trans. Electron Devices, ED-30 (1983) 122
[25] Chao, C.C. and White, M.H., Solid State Electr. 30 (1987) 307
[26] Dellin, T.A. and McWhorter, P.J., Ext. Abstr. Fall Meeting ECS 1986, Abstract 367
[27] Kapoor, V.J. and Bibyk, S.B., *Physics of MOS insulators*, Eds. Lucovsky, G., Pantelides, S. and Galeener G. (1980) 117
[28] Libsch, F.R., Roy, A. and White, M.H., 8th NVSM, Vail (1986)
[29] Cricchi, J.R., Fitzpatrick, M.D., Blaha, F.C. and Ahlport, B.T., IEEE Trans. Nucl. Sci NS-24 (1977) 2185
[30] Peckerar, M.C. and Bluzer, N., IEEE Trans. Nucl. Sci. NS-27 (1980) 1753
[31] Vail, P., Proc. ECS Vol 83-8 (1983) 207
[32] Knoll, M.G., Dellin, T.A. and Jones, R.V., IEEE Trans. Nucl. Sci. NS-30 (1983) 4224
[33] Brown, W.D., Jones, R.V. and Nasby, R.D., Solid State

Electr. 29 (1985) 877
[34] Khang, D.H., Sze, S.M., Bell Syst. Tech. J., 46 (1967) 1283
[35] Frohman-Bentchkowsky, D., IEEE J. Sol. St. Circ., SC-6 (1971) 301
[36] Anantha, N.G., Dockerty, R.C., Lucy, J.M., IBM Tech. Discl. Bull., 17 (1975) 2311
[37] Lenzlinger, M., Snow, E.H., J. of Appl. Phys., 40 (1969) 278
[38] Klein, R., Owen, W.H., Simko, R.T., Tchon, W.E., Electronics, (Oct. 11, 1979) 111
[39] Landers, G., Electronics, (June 30, 1980) 127
[40] Rossler, B., IEEE Trans. Electron Devices, ED-24 (1977) 806
[41] Guterman, D.G., Rimawi, I., Chiu, T., Halvorson, R., McElroy, D., IEEE Trans. Electron Devices, ED-26 (1979) 576
[42] Maes, H.E., Groeseneken, G., 5th NVSM, Monterey, March 1982
[43] Samachisa, G., Su, C.-S., Kao, Y.-S., Smarandoiu, G., Wong, T., Hu, C., Digest Tech. Papers ISSCC87 (1987) 76
[44] Mukherjee, S., Chang, T., Pang, R., Knecht, M., Hu, D., Tech. Digest IEDM 1985 (1985) 616
[45] Lee, J., Dham, V.K., Tech. Digest IEDM 1983 (1983) 589
[46] Guterman, D., Houch, B., Starnes, L., Yeh, B., Tech. Digest IEDM 1986 (9186) 826
[47] Mielke, N., Fazio, A., Liou, H.-C., presented at IRPS87 (1987)
[48] Maes, H.E., Groeseneken, G., Tech. Digest IEDM 1984 (1984) 476
[49] Wegener, H.A.R., 8th NVSM, Vail, Aug. 1986
[50] Dickson, J.F., IEEE J. Sol. St. Circ., SC-11 (1976) 374
[51] Cuppens, R., Hartgring, C.D., Verwey, J.F., Peek, H.L., Vollebregt, F.A.H., Devens, E.G.M., Sens, I., IEEE J. Sol. St. Circ., SC-20 (1985) 603
[52] Yoshikawa, R., Matsukawa, N., 8th NVSM, Vail, Aug. 1986
[53] Mielke, N., Purvis, L., Wegener, H.A.R., 8th NVSM, Vail, 1986
[54] Lai, S.K., Lee, J., Dham, V.K., Tech. Digest IEDM 1983 (1983) 190
[55] Jenq, C.S., Chiu, T., Joshi, B., Hu, J., Tech Digest IEDM 1982 (1982) 309
[56] Sarma, K.C., Owens, A.H., Pan, D.S., Rosier, B.K., Yeh, L., 8th NVSM, Vail, Aug. 1986
[57] Kamiya, M., Kojima, Y., Kato, Y., Tanaka, K., Hayashi, Y., 1982 IEDM Tech. Digest (1982) 741
[58] Mizutani, Y., Makita, K., 1985 IEDM Tech. Digest (1985), 63
[59] Wu, A.T., Chan, T.Y., Ko, P.K., Hu, C., 1986 IEDM Tech. Digest (1986), 584
[60] DiMaria, D.J., De Meyer, K.M., Serrano, C.M., Dong, D.W., J. Appl. Phys. 52 (1981) 4825
[61] Ito, T., Hijiya, S., Nozaki, T., Arakawa, H., Ishikawa, H., Shinoda, M., IEEE Trans. Electron Devices ED-26 (1979) 906
[62] Matsukawa, N., Morita, S., Nozawa, H., Ext. Abstr. Int. Conf. on Sol. St. Dev. and Mat. (1984) 261
[63] Nozawa, H., Matsukawa, N., Morita, S., IEEE Trans. Electron Devices, ED-33 (1986) 275
[64] Stewart, R.G., Ipri, A.C., Faraone, L., Cartwright, J., Schlesier, K., 1984 IEDM Tech. Digest (1984) 472
[65] Chan, T.Y., Young, K.K., Hu, C., IEEE Electr. Dev. Lett., EDL-8 (1987) 93
[66] Lundkvist, L., Lundstrom, I. and Svensson, C., Solid State Electr. 16 (1973) 811
[67] Maes, H.E. and Van Overstraeten, R., J. Appl. Phys. 47 (1976) 667
[68] Heyns, G. and Maes, H.E., *Insulating films on Semiconductors 86*, Eds. De Keersmaecker R. and Declerck, G., (1986) to be published
[69] Williams, R.A. and Beguwala, M.E., Trans. Electron Devices, ED-25 (1978) 1019
[70] Maes, H.E. and Usmani, S., J. Appl. Phys. 53 (1981) 7106
[71] Suzuki, E. and Hayashi, Y., J. Appl. Phys. 52 (1981) 6377
[72] Heyns, G., Ph. D. Thesis, K.U. Leuven, Belgium 1986, unpublished
[73] Maes, H.E. and Vandekerckhove, E., Ext. Abstr. Fall Meeting ECS 1986, Abstract 369
[74] Roy, A., Libsch, F.R. and White, M.H., Ext. Abstr. Fall Meeting ECS 1986, Abstract 370
[75] Caywood, J.M., Prickett, B.L., Xicor Inc. Report
[76] Shiner, R.E., Mielke, N., Haq, R., Tech. Digest IRPS 1983 (1983) 248
[77] Lai, S.K., Dham, V.K., Guterman, D., Tech. Digest IEDM 1986 (1986) 580
[78] Nakamura, H., Sawase, T., Kihara, T., Matsubara, K., Digest Tech. Papers ISSCC87 (1987) 194

Session A3.1

MOS Memories

Chairman: G. Declerck

Wednesday, September 16, 1987

COMPARISON OF 11.5 μm² STACKED CAPACITOR AND TRENCH CAPACITOR CELLS IN A 4MBIT TEST MEMORY

L.Risch, W.Sesselmann and R.Tielert
Siemens AG, Central Research and Development,ME23
Otto Hahn Ring 6, 8 Munich 83, FRG

A 1Mbit DRAM testcircuit with stacked capacitor or alternatively with trench capacitor cells has been fabricated in 4Mbit design rules and technology. The smaller capacitances obtained by the stacked capacitor at same cell sizes is compensated by reduced leakage currents and α-particle sensitivity. Both cells need a cell voltage of 5V for reliable operation.

1.INTRODUCTION

The stacked capacitor cell (STC) as well as trench capacitor cell are considered the most promising candidates for the 4Mbit DRAM generation. Both cells have specific advantages and disadvantages. Yet, a comparison of those cells in a geometry typical of a 4Mbit DRAM has not been reported. For such purpose we developed a complex test memory with 2.4μm x 2.4μm wordline and bitline pitch. Figures of merit measured on those cells will be reported in this paper.

2.TECHNOLOGY

The 1Mbit DRAM, displayed in Fig.1, has been fabricated in a 1μm CMOS n-well process using polycide for the wordline (running vertically) and Al1 for the bitline (running horizontally).



Fig.1 1Mb-test circuit with external timing

Al2 is used in the periphery and in several memory blocks for the realization of a divided bitline architecture. For the STC cell a third level of poly-Si has been added. Layouts for both cells are shown in Fig.2.

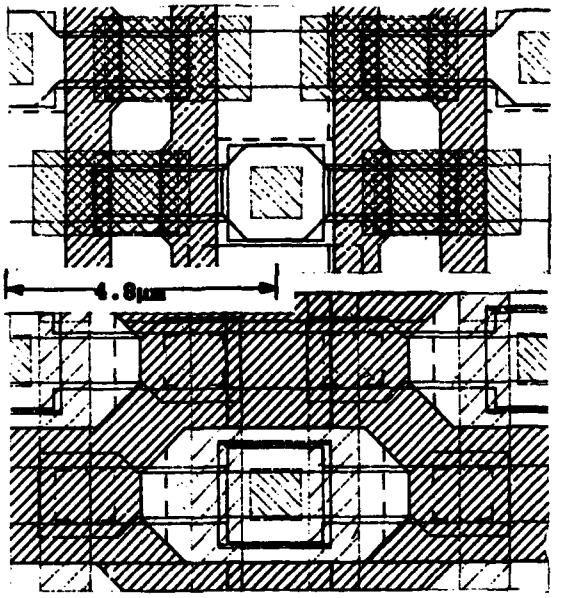


Fig.2 Layout of STC-cell(top) and trench cell(bottom)

3.1 TRENCH CELL

The depletion doped trench cell has been described in /1/. Trench depth

is 4.5 µm, the lateral size is 0.9µm x 1.0µm. A shallow n-layer with a doping concentration of 5*10E18 cm-3 at the sidewalls of the trench is obtained by doping with As. SiO2 with a thickness of 13nm is grown as dielectric. A p-well with a doping concentration of 1.5*10E16 cm-3 is used in order to suppress punch through from trench to trench at a spacing of 1.8µm.

3.2 STACKED CAPACITOR CELL

A cross section of the 11.5µm² cell and the topology in the memory array is shown in Fig.3. The planar area of the bottom electrode and the sidewalls contribute to the cell capacitance. This increase of capacitance of about 50% has been described in /2/. The vertical structure of the STC cell is kept as planar as possible to avoid problems in lithography and etching. Therefore the thickness of the bottom capacitor electrode should be limited to less than 0.4µm.

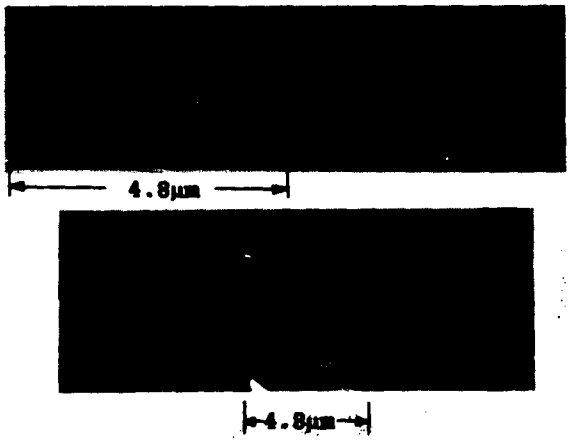


Fig.3 Cross section of the STC-cell and topology of the cell-array

4. EXPERIMENTAL RESULTS

4.1 Dielectric

For a 4Mb STC cell a high-ϵ dielectric must be used to get sufficient cell capacitance. Leakage currents through Si3N4 layers (ϵ≈6.5) with a thickness of 15nm and 20nm show ohmic conduction at low and Frenkel-Poole conduction at high gate voltages, see Fig.4. Operating at ½Vcc cell plate

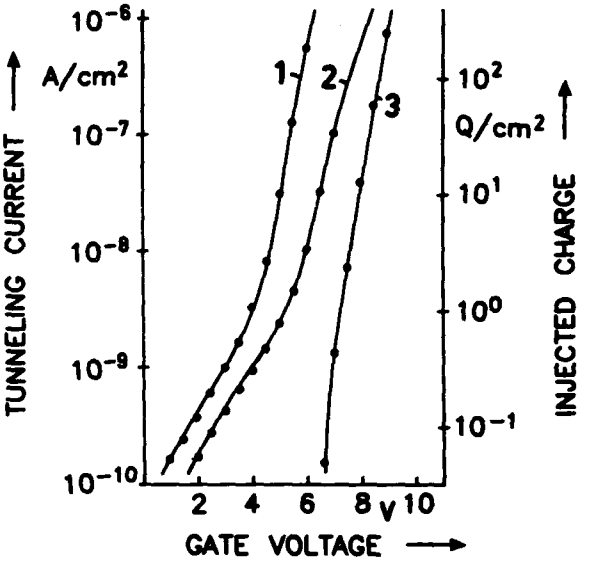


Fig.4 Tunnelling currents of Si3N4 15nm(1),20nm(2) and SiO2-trench 13nm(3)

voltage the injected charge is mainly due to ohmic currents and decreases only linearly for thicker dielectrics. Therefore 15nm of Si3N4 has been used. From reliability measurements the maximum allowed charge injected into this dielectric has been estimated to 1Q/cm². This corresponds to a lifetime of the capacitor of more than 10 years.

The thermally grown SiO2 dielectric with a thickness of 13nm used in the trench cell has orders of magnitude lower leakage currents. Lifetime of the dielectric is in this case not primarily limited by the average injected charge, but by weak spots and sharp edges due to the trench geometry.

4.2 Cell capacitance

Cell capacitance as a function of the gate voltage of the cell transistor at typical bias conditions is shown in Fig.5. For both cell types approximately the same threshold voltage and body effect of the transistor is observed. A boosted wordline higher than 7V must be used for write and read cycles. The capacitance of the STC cell depends less on cell plate voltage due to the higher doping concentration of the poly-Si electrode compared to the n- layer. Fig.6 shows the measured mean values of the cell capacitances as a function of wordline and bitline pitch. At a typical 4M pitch of 2.4 um the trench cell yields higher capacitances of 10 to 20%.

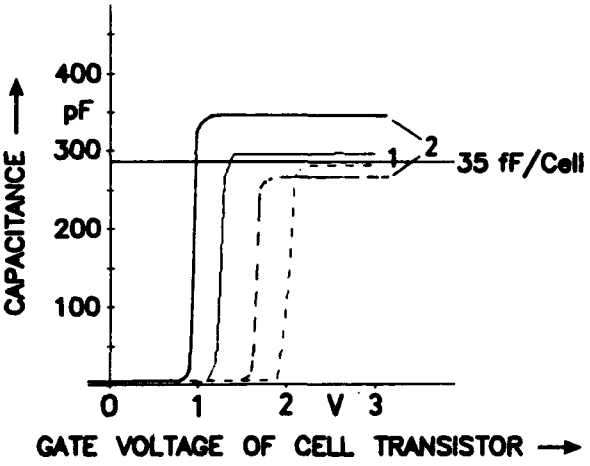


Fig.5 Capacitances of 8k-cells (stacked 1,trench 2) with typical bias.Vbs=-2.5,Vbl=5V (turn off), Vbs=-7.5V,Vbl=0.1V (write)

4.3 Substrate leakage currents

Due to the wide separation of the junctions in the STC cell there is no need for a high doping of the substrate or a p-well to suppress punch through. The low leakage current of a 8k STC array with junction areas below $1um^2$ is shown in Fig.7. Junction breakdown to the 5Ωcm substrate occurs at bitline voltages exeeding 12V.

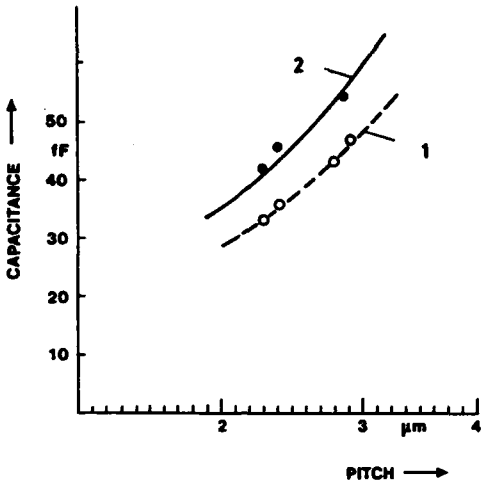


Fig.6 Capacitances of STC-cell(1) and trench cell(2) as function of word-and bitline pitch

For trench cells the junctions are very close and the area of the junction is increased by a factor of 20. Hence, punch through has to be accounted for by adequate doping of the substrate. Limits are given by avalanche or Zener breakdown of the n-layer and by the tolerable body effect of the cell transistor. These effects result in increased generation currents at low bitline voltages and decreased breakdown voltage of 9V , see Fig.7.

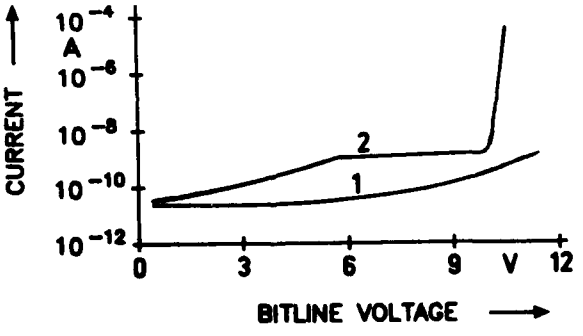


Fig.7 Leakage current of 8k-cells as a function of bitline voltage. (stacked 1,trench 2),Vwl=7.5V,Vpl=2.5V

4.4 α-sensitivity

The small α-particle sensitivity of the STC cell has been confirmed. Compared to the trench cell at a pitch of 2.4µm the STC cell collects a maximum charge of 35fQ and the trench cell 60fQ from an α-particle track which generates a total charge of 220fQ (Ra226), see Fig.8. From these collected charges the soft error rate of the DRAM can be predicted /3/ to be smaller than 500 FITS for the given cell capacitances.

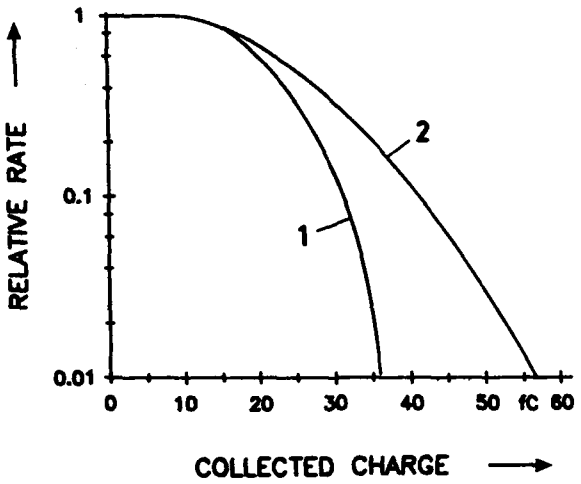


Fig.8 Relative rate of collected charge from an α-particle track. STC-cell(1),trench cell(2)

4.5 Signal to noise margin

In a 1T-cell the signal being detected by the sense amplifiers depends on the ratio of cell to bitline capacitance. With 128 cells per bitline a bitline capacitance of 240fF has been measured for the STC cell and 310fF for the trench cell. Assuming a tolerance of the cell capacitance of 20% and a 4.5V power supply this corresponds to an ideal sense signal of 480mV and 470mV respectively. Taking into account a minimum sense signal of 30mV, midlevel tolerance of 20mV, bitline coupling of 15mV, restore "1","0" of 20mV,10mV, Vt sense amplifier of 30mV and leakage currents of 20, 10mV (Fig.7), an α-particle margin of 40 fQ for the STC cell and 60fQ for the trench cell is estimated for reliable operation.

5.CONCLUSION

Operation of the STC cells as well as of trench cells has been demonstrated in an experimental DRAM in 4Mbit layout rules. Critical process features for the STC cell are patterning of the bottom capacitor electrode and the topology of the cell. Also the long term reliability of the new Si3N4 dielectric still has to be verified on production level. The trench cell needs a field isolation with minimum dimensions and more complicated etching and doping techniques.
Due to the smaller leakage currents, smaller α-particle sensitivity and smaller bitline capacitances a minimum cell capacitance of 25 to 30fF is estimated for the STC cell whereas the trench cell requires a cell capacitance of 30 to 35fF for save operation.

Acknowledgement

This report is based on a project which has been supported by the Minister of the FRG under the support-no NT 2696. For the contents the authors are responsible alone.

References

/1/ K.H.Küsters et al.,A High Density 4Mbit DRAM Process Using a Fully Overlapping Bitline Contact (Fobic) Trench Cell, 1987 Symposium On VLSI TECHNOLOGY, Karuizawa, Japan
/2/ K.Lau et al.,Stacked Capacitors for Mb DRAM, ESSDERC 86,pp.191-192
/3/ W.Krautschneider et al.,Prediction of Soft-Error Rate of 4M-DRAM, this conference.

Isolation - Related Leakage in a 4 Mb DRAM Cell

P. A. Murkin, H.- M. Muehlhoff, S. Roehl, W. Meyberg,
W. Mueller, W. Bergner and R. Kircher

Corporate Research and Technology,
Technology Centre for Microelectronics,
Siemens AG,
Otto-Hahn-Ring 6, 8000 Munich 83, West Germany.

The isolation-related leakage occurring within a 4 Mb DRAM cell containing a depletion-type trench capacitor is discussed. The main isolation-dependent leakage paths discussed are trench-trench punchthrough, inter-bit-line leakage via punchthrough or parasitic transistor action under the word-line, and leakage between the transfer gate and the trench capacitor. Measurements of specific test structures and also of an 8K cell array confirming good control of these leakage mechanisms are presented.

1. INTRODUCTION

The achievement of low levels of parasitic leakage is increasingly important in Mb DRAMs, as critical isolation widths are shrunk to sub-micron dimensions. In this paper measurements of isolation-related leakage currents within the cell area of a 4 Mb DRAM are presented. The different parasitic leakage paths are described and their relative importance is shown using experimental results. The total leakage current measured on an 8K cell array confirms satisfactory control of these effects.

2. DESCRIPTION OF 4 MB DRAM

The submicron CMOS technology used for the 4 Mb DRAM has 0.9 micron design rules and is configured with triple level polysilicon and single level metallisation. The cell capacitor is a depletion-type trench capacitor with an effective dielectric thickness of 13 nm and a cell capacity of 40 fF. The first layer of polysilicon is used as the cell electrode, the second for the word-line (shunted with metal to lower the resistance) while the third layer is a low resistivity polycide which is used for the bit-line. A self-aligned bit-line contact is used to achieve a cell size of 2.3 x 4.6 = 10.6 sq. microns (1).

3. DESCRIPTION OF LEAKAGE PATHS

A schematic layout of a portion of the cell area is shown in fig.1. The labelled leakage paths between the features are:
(A) trench-trench punchthrough; (B) inter-bit-line leakage under the word-line in the region of the self-aligned contact; (C) inter-bit-line leakage under the word-line; (D) leakage between the transfer gate and the trench capacitor. Cross-sections through these regions are also shown in the figure. The isolation is achieved by using a modified local oxidation process.

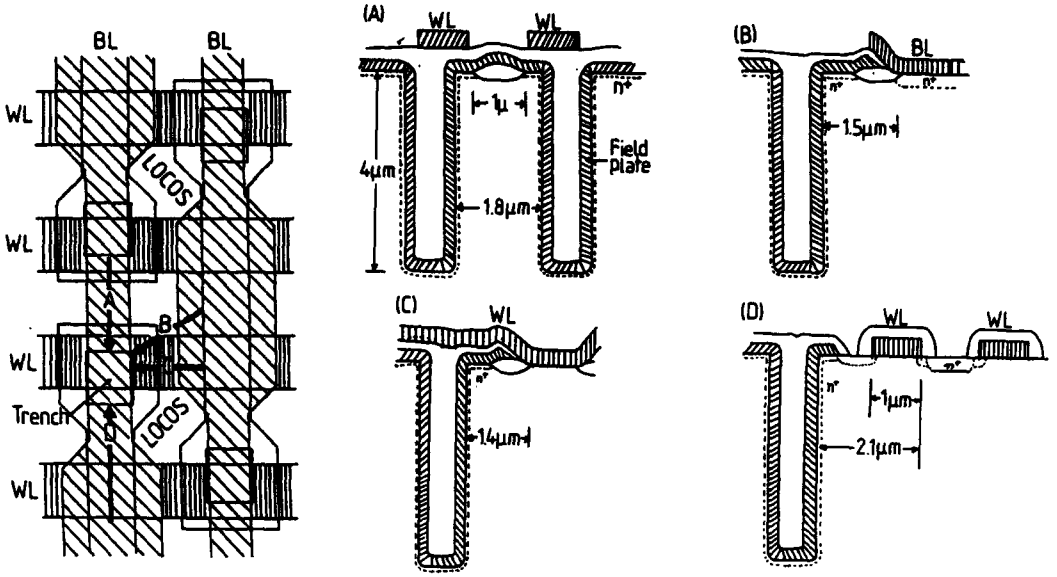


Fig.1. Schematic layout of a portion of the cell area showing isolation-related leakage paths and also cross-sections in these regions.

4. RESULTS

4.1 Trench-trench leakage

The leakage between two adjacent trench capacitors is controlled by the use of a p well which is doped to at least 10^{16} cm^{-3} down to the bottom of the trenches: this doping criterion is derived from device simulations. To ensure that no punchthrough occurs, good control is required of factors which affect the trench-trench spacing, e.g. the etched profile of the trenches and the diffusion of the As-doped HiC varactor doping. Fig. 2 shows trench profiles obtained using different etch techniques : for the same nominal spacing of 1.8 microns, the actual minimum spacing is smaller by 0.7 and 0.25 microns respectively.



Fig.2. Trench profiles obtained using different etch processes. (The nominal spacing of the trenches is the same.)

Fig. 3 shows the results of trench-trench punchthrough measurements : for the nominal separations shown, only those wafers with the larger decrease in spacing show punchthrough. Results from numerical simulations utilizing a newly developed 3D device simulatôr (2) show that for actual trench spacings greater than one micron punchthrough is not seen below 10V, which is confirmed by measurements.

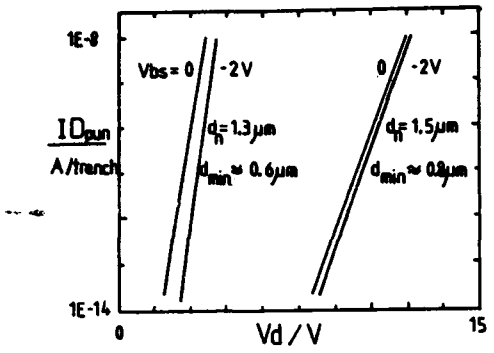


Fig.3. Measurements of trench-trench punchthrough leakage for various trench separations.

4.2 Inter-bit-line leakage

Leakage between the bit-lines occurs between the trench and the adjacent n+ regions via paths determined by the word-line biases. There are two components of this leakage : simple punchthrough and also a parasitic transistor lying under the polysilicon word-line. Direction (C) represents the latter parasitic transistor whilst (B) is a combination of both of these mechanisms.

Fig. 4 shows that the characteristic of the parasitic transistor (C) is strongly dependent upon the substrate-bias applied. With the use of substrate-bias satisfactory performance is obtained, even under the worst-case operating condition of Vbs=-1V, 90C.

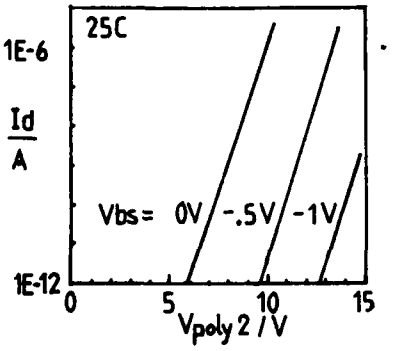


Fig.4. Measured characteristics of the parasitic transistor under the word-line (C).

The formation of the self-aligned contact degrades the isolation between bit-lines. This is due to the field oxide under the word-line being further etched-back and also that another parasitic field-plate is available. Fig. 5 shows the punchthrough performance between bit-lines both within the vicinity of the self-aligned contact and outside of it. It is clear that degradation of the isolation behaviour does occur, but that under the operating conditions the use of substrate-bias prevents punchthrough.

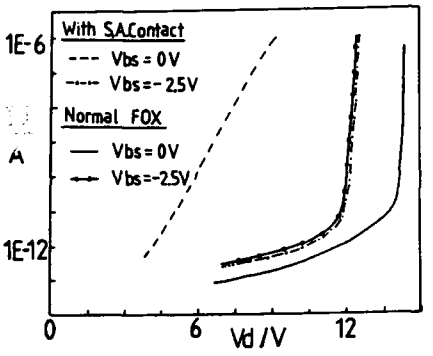


Fig.5. Comparison of punchthrough performance between the bit-lines both in the vicinity of the self-aligned contact and outside of it.

4.3 Transfer-gate leakage

The subthreshold characteristics of the transfer-gate transistor are shown in Fig. 6 for the worst-case operating condition of Vbs=-1V and 90C. Results from the 2D simulator GALENE (3) show that for trench to transfer-gate spacings of smaller than 0.3 microns no punchthrough leakage occurs because of the high p well doping used to prevent trench-trench punchthrough.

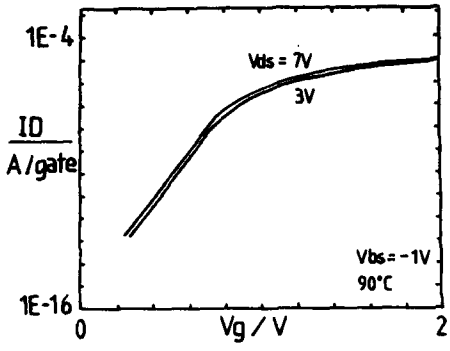


Fig.6. Subthreshold characteristic of the transfer-gate.

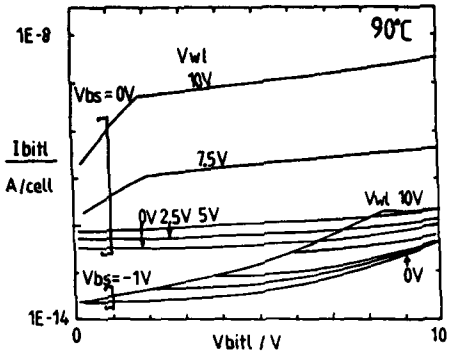


Fig.8. Cell array measurements with inter-bit-line leakage.

4.4 8K Cell Array

Measurements on an 8K cell array at 90C are shown in Figs. 7 and 8. In Fig. 7 both the bit-lines were swept together at the same potential, with the two word-lines also tied to one another and swept over the same potential range, thus enabling a measurement of the substrate leakage.

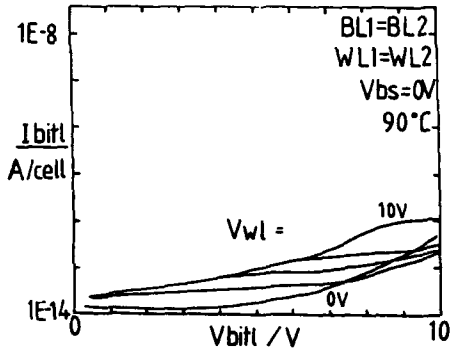


Fig.7. Cell array measurements with the bit-lines tied together.

In Fig. 8 one bit-line was held low and the other was ramped up, so that inter-bit-line leakage was possible. One word-line was held at 2.5V while the voltage applied to the other was varied, thus enabling a separation of the leakage paths (B) and (C). At zero substrate-bias the parasitic field oxide transistor (C) turns on, but this can be suppressed by -1V reverse bias. From these measurements it is confirmed that the total leakage is less than 1pA per cell at 90C at a bit-line bias of 7.5V and word-line voltages of 10V, for operation with -2.5V substrate bias.

5. CONCLUSIONS

The leakage current paths related to the device isolation in a 4 Mb DRAM cell have been described and experimental results showing control of the isolation characteristics presented. Measurements of the leakage within an 8K cell array are also reported and show inter-bit-line leakage to be very low.

REFERENCES

(1)Kuesters, K.- H. et al, Proc. of VLSI Symposium 1987, pp.93,94.

(2)Bergner, W. and Kircher, R., SITAR: An efficient 3D device simulator for trench cells; to be published

(3)Engl, W.L. and Meinerzhagen, B., Proc.IEEE, Vol.71(1983), p.10

A 18 μm^2 cell for Megabit CMOS EPROM

E. Camerlenghi, P. Caprara and G. Crisenza
SGS Microelettronica
via C.Olivetti 2
20041 Agrate Brianza (MI)-Italy

1. INTRODUCTION

Due to recent developments in computers, non volatile memories have been taking on a greater importance in the semiconductor market, and consequently higher density and higher performance memories are required. Regarding the EPROM (electrically programmable read only memories) area, three qualities are required to define a good EPROM technology [1] [2]:

Compactness, reflecting on the possibility of reducing chip dimensions and costs;

Speed, reflecting on better device performances, reducing access, programming and testing times;

Manufacturability; with this expression we indicate an overall valuation of the critical extent of the architecture, that is to say the sensitivity of that architecture to variations in the process technology parameters.

This aspects are underlined in this paper in which we present the structure and the full electrical characterization of the EPROM memory cell, extensively tested on a 64 Kbit device and currently used in our line for a compact and fast 1 Mbit EPROM memory. The cell area is 18 μm^2, which represents one of the most compact EPROM technology for 1Mbit generation devices [3] [4].

2. PROCESS ARCHITECTURE AND CELL STRUCTURE

Many factors contribute to obtaining a good EPROM technology: some of them are strictly related to the cell structure but others are related to the whole process technology and to the particular device organization and design, and for this reason they can't be ignored and are here succinctly reported.

2.1. Process architecture

Our EPROM memory cell is fully compatible with 1 μm n-well CMOS process technology using tantalum disilicide on the gates to lower their resistance and consequently the word line delay (Fig. 1).

LDD structure on n-channel transistors permits enhanced breakdown voltage and consequently internal high bias can be generated. This way, writing times per bit can be reduced.

Shallow junctions are used to enhance architecture compactness and to improve the device speed characteristics reducing source and drain parasitic capacitance. In addition, barrier metal that has been introduced to contact shallow junctions, reduces contact resistivity of the n-type contacts (representing, in an n-channel EPROM device, the prevailing contact type) and improves reliability characteristics.

In Tab 1 the main characteristics of this technology are summarized.

Minimum feature size	$1\mu m$
Technology	Double poly, n-well, CMOS
n-ch.transistor L_{poly}	$1.2\mu m$
p-ch.transistor L_{poly}	$1.6\mu m$
n-ch. threshold voltage	0.75 V
p-ch. threshold voltage	0.95 V
n-ch. breakdown BV_{DSS}	17 V
p-ch. breakdown BV_{DSS}	13 V
1 Mbit die size	71 $Kmil^2$

TABLE 1: Main parameters characterizing the process technology

2.2. Cell structure

In Fig. 2 and 3 the layout of the transistor memory cell and a SEM picture of the actual processed cell inside the device matrix are reported.

As it can be noted in Fig. 2, the memory cell architecture is a conventional one, characterized by a symmetrical layout.

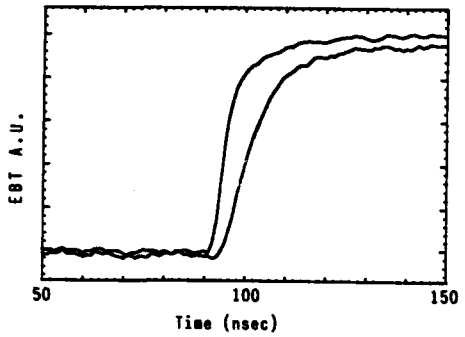


FIGURE 1: EBT data showing the signal propagation delay on a word line

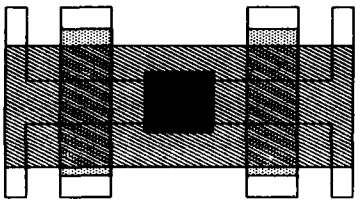


FIGURE 2: Cell layout



FIGURE 3: SEM top view of the matrix of cells

The memory cell is a self aligned stacked structure with poly-Si and polycide for the floating and control gates respectively as is shown in the cross section reported in Fig. 4. First gate oxide is 28 nm. The interpoly dielectric is a triple layer using SiO_2, Si_3N_4, SiO_2 (O.N.O.), whose equivalent oxide thickness is 30 nm. The cell effective length is 1 μm.

The particular kind of junctions oxidation allows very low bending of the floating and control gates. This way coupling ratios are maintained as closely as possible to the theoretical ones, consequently the cell writing efficiency is maximized.



FIGURE 4: SEM cross section of the memory cell

3. ELECTRICAL RESULTS

In this part we are going to analyze cell behavior and its parasitic effects in relation to variations in the main process parameters. In particular, the channel effective length (L_{eff}) and the drain voltage used to write the cell (Vd_w) are the two quantities most affecting cell characteristics, because of the exponential dependence of the gate and substrate currents on these parameters [5] [6].

In Fig. 5 writing curves are shown for several EPROM cells with different channel effective lengths, in the range between 0.5 and 1.5 μm. More than a 3 V threshold shift is performed in less than 100 μsec by most cells; only the longest of them, whose effective length is 50 % bigger than the project one, requires 200 μsec to perform a 2 V threshold shift.

In curve A of Fig. 6 is reported the locus of L_{eff} and Vd_w which allow a 3 V threshold shift with a writing time of 130 μsec and a gate applied voltage of 12.5 V. Similar curves can obviously be obtained for a different choice of parameters, but these values may be considered a good limit for distinguishing between adequate and bad writing cell performance, for 1Mbit generation device.

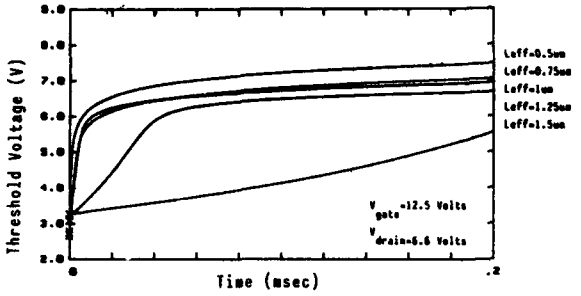


FIGURE 5: Writing curves of different channel effective length memory cells

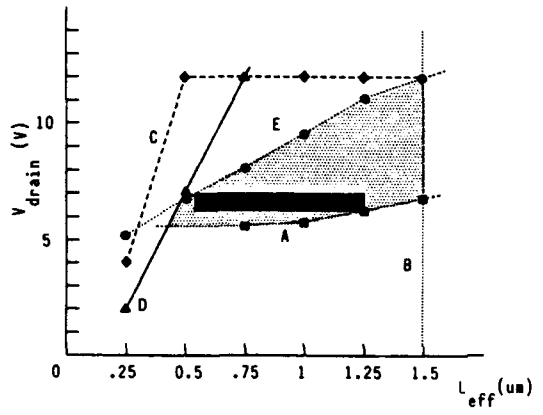


FIGURE 6: Memory cell working area

A : *writing curve* D : *drain turn-on curve*

B : *reading curve* E : *snap-back curve*

C : *breakdown curve*

Reading characteristics for different effective length virgin cells are reported in Fig. 7, in which the worst case reading bias is considered. It can be noted that sufficient drain reading current is measured up to 1.5 μm effective channel length, therefore this value becomes a maximum limit for L_{eff}, causing curve B of Fig. 6.

Grounded gate breakdown voltage for different L_{eff} values has been measured on equivalent transistors having the same device parameters as those of the memory transistor. Data are reported in curve C of Fig. 6. Actually, the grounded gate breakdown isn't the limiting effect to the cell performance; in fact,

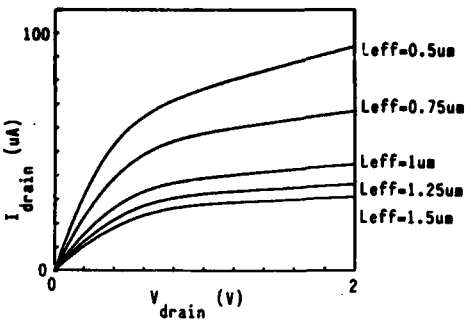


FIGURE 7: Readind characteristics for different L_{eff} cells; Vd=1V; Vg=4V

during the real device operation, caused by the drain capacitive coupling, the cell floating gate voltage assumes values higher than zero, even when the control gate is grounded.

In this situation the so called *drain turn-on* effect may take place, in which leakage of the unselected memory cells connected to the selected bit line can cause a voltage drop of the bit line and consequent degradation of the programming characteristics. In 1Mbit memory device these effects are avoided if less than 100 nAmp leakage per cell is assured, with the gate bias equal to the worst case for the logical zero (that is assumed to be 0.3 V). In curve D of Fig. 6, the locus of L_{eff} and Vd_w satisfying this requirement is reported, again introducing a boundary between good and bad cell performances.

The *snap-back* phenomenon must also be avoided because the high substrate currents accompanying with it may cause latch-up in CMOS devices, with the consequent risk of device damage. In Fig. 8 Vg and Vd pairs causing snap-back onset are reported for different effective length equivalent transistors; for every L_{eff} value there exists a minimum drain voltage that avoid snap-back, independently of the gate (or floating gate) bias. These points are reported in curve E of Fig. 6.

Considering Fig. 6 as a whole, it can be noted that, in the L_{eff} - Vd_w plain, the five curves define a spatial region (dotted in figure) whose points correspond to EPROM cells with a good overall performance. Manufacturing necessities require that such good behaviour be guaranteed in spite of writing drain

voltage and effective length dispersions. Due to these restrictions, the memory cell working area becomes a rectangle. One is shown in Fig. 6, representing an example of manufacturing suitability of the cell design: for a drain voltage of 6.5 V ± 300 mV the cell performances are well above the acceptable limit over a wide range of cell transistor L_{eff} centered around 0.9 μm.

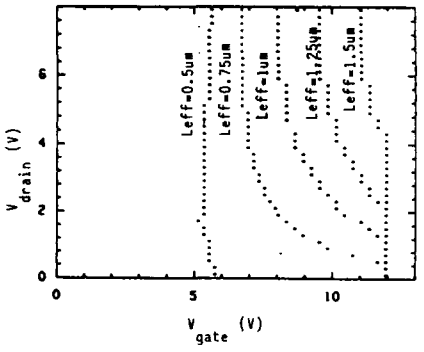


FIGURE 8: Snap-back onset curves for different L_{eff}

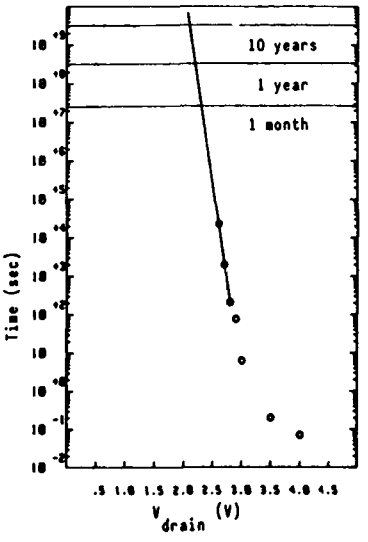


FIGURE 9: Soft-writing curve; the time required to perform 200 mV threshold shift is reported for different drain bias, with 6V applied to the gate

The problems related to the unwanted cell writing that may be introduced by the protracted reading operation (*soft-writing*), have been investigated on the shortest cell, that is the most affected by these effects. Data are reported in Fig. 9 which shows the time required to obtain a 200 mV threshold shift for different drain voltages, with the gate bias at the worst case (6 V) for the reading operation.

Extrapolated data show a safe reading operation for all practical drain bias.

Finally, no charge retention problems have been pointed out by voltage stress measures; furthermore, on 64 Kbit memory devices realized to qualify this technology, temperature tests have been executed and passed.

4. SUMMARY

A process technology for realizing highly compact EPROM devices has been described; in particular the structure and the electrical characterization of a 18 μm^2 memory cell has been reported, pointing out good speed performances together with good manufacturability qualities.

This technology has been verified using 64K memory, and at present it is utilized to manufacture a 1Mbit memory.

ACKNOWLEDGEMENTS

This work has been partially supported by Consiglio Nazionale delle Ricerche, Progetto Finalizzato: "Materiali e Dispositivi per l'Elettronica a Stato Solido"

REFERENCES

[1] J.Lien et al. *IEDM*, (1984) pp. 460-463

[2] K. Yoshikawa et al. *IEDM*, (1984) pp. 456-459

[3] G. Gerosa et al. *IEDM*, (1985) pp. 631-634

[4] K.Komori et al. *IEDM*, (1985) pp. 627-630

[5] C. Hu et al. *IEEE J. Sol. State Circ.*, (1985) vol. sc-20, N.1, 295

[6] B.Eitan et al. *IEEE, Trans. El. Dev.*, (1981) vol. ED-28, N.3 328

A3.1.4

A high performance P-channel EPROM cell

D. Cantarelli, A. Maurelli and L. Baldi
SGS Microelettronica S.p.A.
via C.Olivetti 2 - 20041 Agrate Brianza (MI)-Italy

Recent results show higher gate current in P-channel devices than in N-channel ones under the same bias conditions. By this reason, we produced a P-channel EPROM cell, making use of a standard N-channel cell layout. We present a complete electrical characterization and point out some advantages which make P-channel EPROM cells very promising, especially for future CMOS applications.

1 Introduction

Apart from the pioneering work of Frohman-Bentchkovsky [1], all EPROM cells up to now have been based on N-channel technology. The usual explanation given for this, is that EPROM cells are written via hot electron injection, and hot electrons are easier to generate in N-channel transistors, due to the higher mobility of electrons. A proof of this is the much higher substrate current in N-channel devices, when they are biased for maximum avalanche generation. It has however been recently pointed out [2], [3] that maximum gate currents are much larger in P-channel than in N-channel transistors, given the proper bias conditions. There is therefore no reason why a P-channel EPROM cell could not be produced, and indeed present some advantages over the traditional N-channel cell in terms of writing speed (because of the larger gate current), power consumption (since it is operated at lower gate voltage) and process compatibility (because of the low substrate current). To verify these points, P-channel EPROM cells have been fabricated and the results of their evaluation are presented in this work.

2 Process architecture and cell structure

The process was a P-channel translation of the N-channel process used for a 256Kbit EPROM, having 33 nm gate oxide, 0.4 μm junction depth, self-aligned floating and control gates and ONO as interpoly dielectric. We used the standard N-channel cell layout shown in Fig.1. The memory cell area is 36 μm^2, while the effective length of the transistor is 1 μm and its effective width is 1.4 μm.

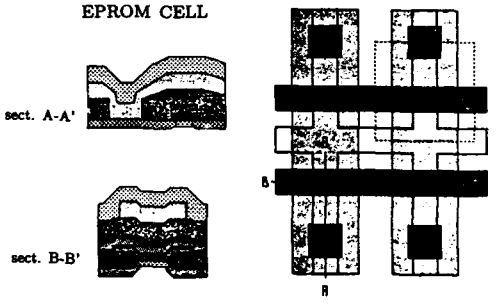


Figure 1: Layout and cross-sections of the EPROM cell

3 Electrical results

Typical transfer characteristics of the P-channel cell after a sequence of 100 μs writing pulses at V_D=-12 V and V_G=-6 V are shown in Fig.2. As it can be seen, very high threshold shifts can be achieved after a rather short writing time. In Fig.3 writing curves are shown for different values of drain voltage and fixed gate voltage. No saturation has been found even increasing the duration of writing pulses up to 0.5 s. In Fig.4 similar curves are shown with gate voltage as

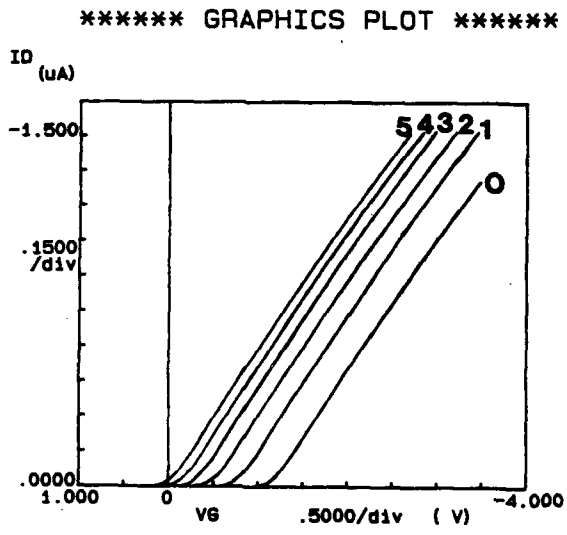


Figure 2: Transfer characteristics of P-channel cell during writing pulse sequence (V_G=-6V, V_D=-12, t_{pulse}=100μs)

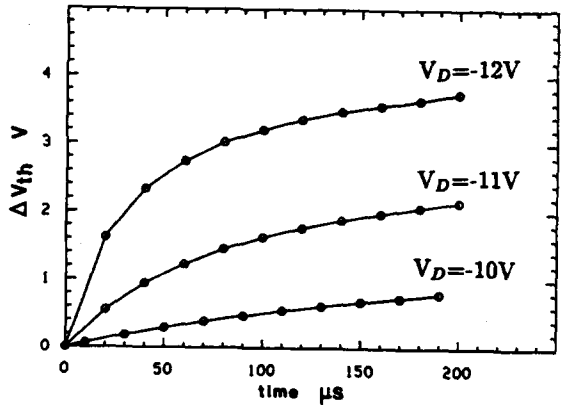


Figure 3: Writing curves - Fixed gate bias (-5V), V_D as parameter

parameter and fixed drain voltage. The maximum of gate current occurs at a gate voltage just above the threshold [3], and therefore, as evident from Fig.4, the lower the gate voltage (absolute value), the higher the threshold shift. This is shown even more clearly in Fig.5 where the threshold shift after a 500 μs writing pulse at V_D=-10 V and V_D=-11 V is give as a function of gate voltage. The peak of threshold shift occurs at V_G=-1.75 V, and corresponds to the maximum of gate current, while at V_G=V_D, which is the best case for

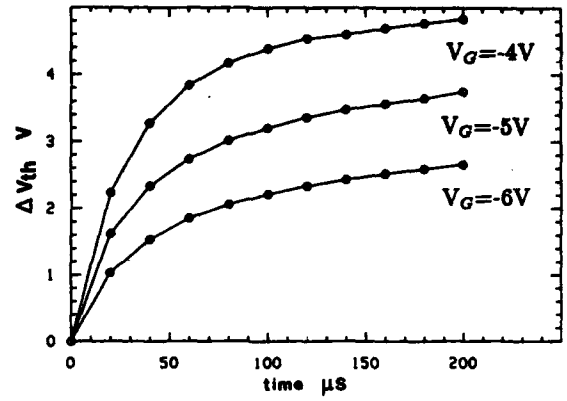


Figure 4: Writing curves - Fixed drain bias (-12V), V_G as parameter

N-channel EPROMs, the writing is negligible. It is evident from the above presented results that different writing mechanisms are active in N-channel and P-channel EPROMs. While in the first ones the energy of the electrons is very important, since they have to be injected against the electric field, in the second ones it does not play any role since the field is always in

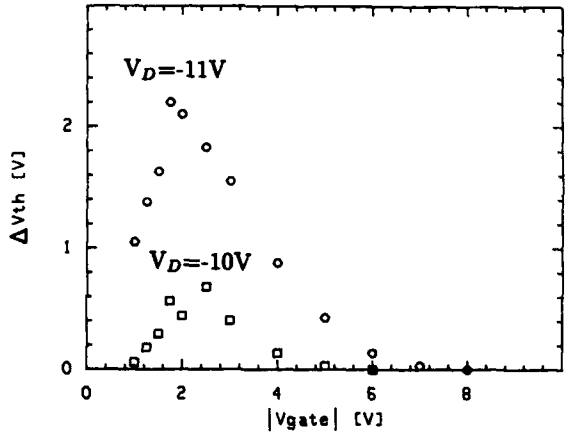


Figure 5: Threshold shift vs gate voltage

favor. This hypothesis is confirmed by curves shown in Fig.6, where the ratio between gate and drain current I_G/I_D is plotted against the substrate current I_B for equivalent N-channel and P-channel transistors. It is evident that while for N-channel cells, the effi-

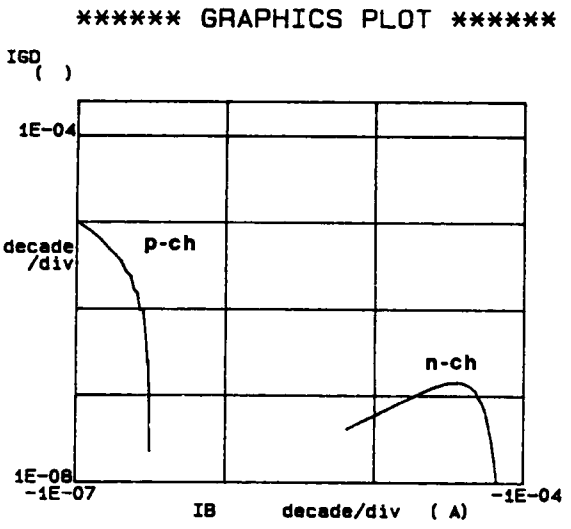


Figure 6: Comparison between N-channel and P-channel equivalent transistors (L=1.5 μm). IGD=I_G/I_D, fixed drain bias ($|V_D|$=9V), 0V$\leq|V_G|\leq$15V.

ciency of gate current generation is very low, asking for high drain and substrate currents, it is very high for P-channel cells. Cancellation and retention performances are comparable with those of N-channel cells and will not be discussed in details. Soft writing, and this is another advantage with respect to N-channel cells, is not a problem at all, because typical reading biases forbid electron injection [3].

4 Conclusions

The results of the preliminary evaluation of a P-channel EPROM cell have shown the feasibility of this device. Some writing characteristics appears to be superior to those of N-channel cells: in particular, the low gate writing voltage allows a reduction in the capacitive coupling between floating and control gate, and therefore a reduction in cell area, while the low drain current during the writing cycle could be compatible with on-chip high voltage generators. Moreover, the very low value of substrate current during writing prevents the risk of latch-up triggering in CMOS devices and makes this kind of cell compatible even with a N-well process. Among the drawbacks, we should mention the lower gain of the P-channel transistor, which affects the reading speed, and the fact that the threshold of the written cell is shifted towards the depletion region, making the cell decoding more complex. However, taking into account that the cell has not been designed to take full advantage of the new technology, it could be concluded that P-channel EPROM cells appear to be promising for future high density applications and well worth investigating.

References

[1] Frohman-Bentchkovsky, *Appl. Phys. Lett.*, 18, 8 Apr. 71, 332

[2] K.K.Ng and G.W.Taylor, *IEEE Trans. Electron Dev.*, vol. ED-30, n.8,p.871, 1983

[3] L.Baldi and M.Tosi, *Proceedings of ESSDERC 1986,* Abstract 86

Session B3.1

Latch-Up

Chairman: M. Montier

Wednesday, September 16, 1987

THE INFLUENCE OF LIFETIME ON THE LATERAL PARASITIC BIPOLAR TRANSISTORS IN CMOS.

L. Deferm, G. Romaen, C. Claeys and R. Mertens

I.M.E.C.
Kapeldreef, 75
B-3030 Leuven - Belgium

The electrical characteristics of lateral bipolar transistors are very important in predicting latch-up. An increase of the current amplification factor is noticed when the high current level injection regime is reached, but this only for devices with a relative low base carrier lifetime. For low minority carrier lifetimes in the base the beta versus collector curve will show a bump.

1. INTRODUCTION

Scaling down CMOS technology leads to a higher sensitivity for latch-up occurence. This phenomenon is strongly related to the behaviour of the parasitic bipolar devices. In a pwell process the vertical bipolar transistors are NPN and the lateral are PNP devices. The DC-characteristics of the vertical transistors can be very well defined from their implantation profiles, unless the surface or space-charge recombination current is very high. On the contrary, the electrical behaviour of the lateral pnp , formed by a p+ diode as emitter, the substrate as base and the pwell as collector, is more difficult to predict Due to a field implantation profile a vertical field pushes the holes, which are injected in the base, away from the surface, resulting in a nonuniform injection along the emitter edge and a lower current amplification factor. Also the effective volume of the base, which influences the beta of the bipolar transistor, is depending on the lateral and vertical dimensions, the implantation profiles and the lifetime of the minority carriers. Especially the evaluation of the minority carrier lifetime and the way in which the current flows make it difficult to predict the behaviour of a parasitic bipolar pnp in a non-epi substrate.

2. APPROXIMATE CALCULATIONS OF THE COLLECTOR CURRENT, THE BASE RECOMBINATION CURRENT AND THE BASE INJECTION INTO THE EMITTER CURRENT

The calculation of the current densities are performed for an uniform doped base, with a lateral width of W_b, a n-type base concentration equal to N_b and a p-type emitter concentration of N_e. The influence of the surface recombination current can be neglected due to an electrical field which is for the minority carriers directed towards the bulk. This field is the result of a field implantation at the surface of the base. In order to calculate the different current densities it is necessary to solve the current and continuity equations. The following one dimensional formulas are used:

$$J_p = - q D_p \partial p(x)/\partial x + q \mu_p p(x) E(x)$$

$$\partial J_p/\partial x + q \partial p(x)/\partial t + q r = 0$$

$$\partial p(x)/\partial t = 0 \text{ in steady state regime}$$

D_p and μ_p are respectively the diffusion constant and the mobility for holes, p(x) is the hole concentration distribution, Jp is the hole current density and E is the electrical field in the base, caused by the distribution of the majority carrier concentration, which becomes important at high level injection. This electrical field is given by:

$$E = -(k\,T / q)*(1/n(x))*\partial n(x)/\partial x$$

In the case of neutrality n(x) can be replaced by (Nb + p'), where p' is the excess hole concentration in the base. The hole current density can be expressed as follows:

$$J_p = -q\,D_p\,(N_b+p_0+2p')/(N_b+p')\;\partial p'(x)/\partial x$$
$$J_p = -q\,D_p\,C(x)\;\partial p'(x)/\partial x$$

p_0 is the hole concentration at the emitter edge with a Vbe of 0 V.
The effective hole diffusion constant is equal to Dp*C(x) [1]. This value is current and position dependent. An approximation can be made in which p'(x) in C(x) is replaced by p' at the emitter edge as done by Chou [2]. Solving the continuity equation results in the following formulas for the current densities, which are valid in high and low current injection regime.

$$Jc = q\,D_p\,\sqrt{C}\;p'(\text{at emitter edge})/(L_p\,F)$$
$$F = \sinh(W_b/(\sqrt{C}\,L_p))$$
$$L_p = \sqrt{(D_p\,\tau)}\quad \text{hole diffusion length}$$
τ = minority carrier lifetime
Jc is the collector current density

Figure 1 shows the influence of C on the collector current, where the curve with the label 'no field' is the collector current density, where C is equal to 1 for the whole Vbe range. Also the ratio between the two curves is shown, indicating no difference at low level injection and a factor of two at high level injection.

The base recombination current density is given by the next formula, where W is either the lateral base width Wb or the wafer thickness, if the bulk recombination current density is desired.

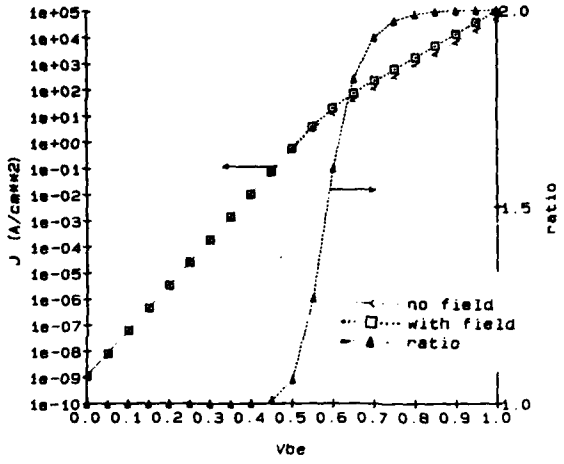


FIGURE 1: Collector current density with and without the field and the ratio of both.

$$J\ \text{rec.} = q\,J_p\,\sqrt{C}\;F1\;p'(\text{at emitter edge})$$
$$F1 = (\cosh(W/(\sqrt{C}\,L_p)) - 1) / F2$$
$$F2 = Lp\,\sinh(W/(\sqrt{C}\,L_p))$$

For both large and small values of W, the base recombination current density is depending on the lifetime: increasing lifetime is decreasing current.
The lifetime is assumed to be independent of the current, but in reality it increases at high current level injection [3]. For small values of the base width compared to the diffusion length the ratio between collector and base recombination current is depending on lifetime and on the value of Vbe (figure 2).

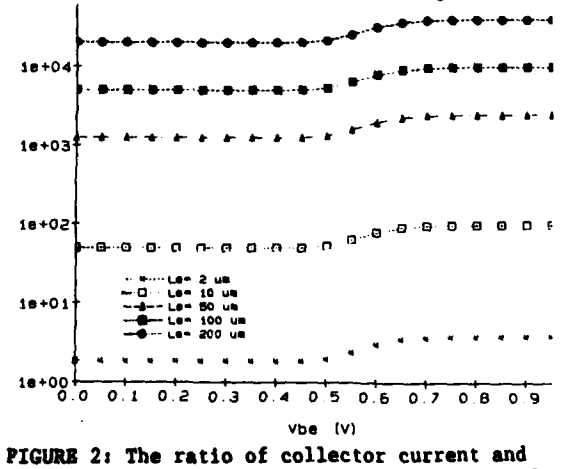


FIGURE 2: The ratio of collector current and base recombination current density for different diffusion lengths (W_b= 2µm).

An increase of this ratio at high level injection is noticed due to an increase in the collector current density. The influence of this ratio on the real current amplification (beta) is depending on the value of the base-into-emitter injection current compared to the value of the base recombination current.

$$Jemitter = q\, D_n\, n'_{(at\ emitter\ edge)} / We$$

We is the width of the emitter

It is assumed that the emitter current is not at high level injection, due to the high concentration in the emitter. For the calculation bandgapnarrowing effects described by Slotboom and de Graaff [4] are included. If the base recombination current is dominant compared to the base-into-emitter injection current until very high level injection, a bump in the beta curve is noticed (figure 3).

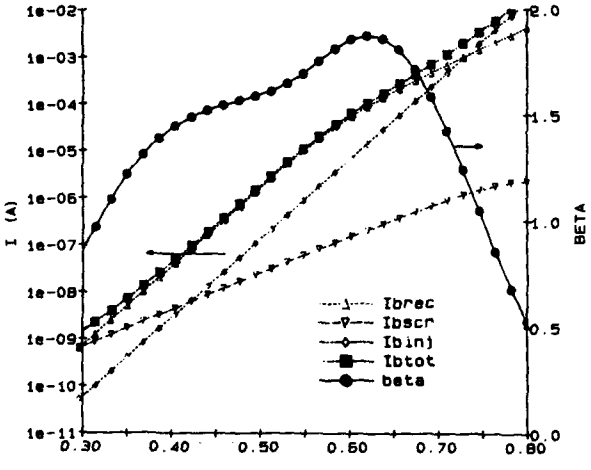


FIGURE 3: Base current components and beta for a lateral pnp with τ= 100 ns. Ibrec is the baserecombination current, Ibscr is the recombination in the space charge layer, Ibinj is the injection into the emitter and Ibtot is the total base current.

Also two dimensional simulations, using the program PISCES, show a bump in the current amplification curve for a minority carrier lifetime smaller than 5 µsec. At higher values of the lifetime this bump disappears. Only two values for the lifetime are presented in figure 4.

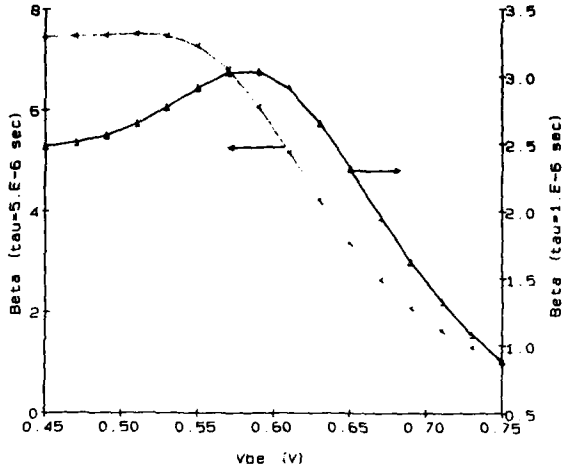


FIGURE 4: Simulated beta curves for two different lifetimes.

3. EXPERIMENTAL DETERMINATION OF LIFETIME IN THE BASE REGION OF PARASITIC PNP TRANSISTORS.

Recombination lifetime of the minority carriers in the substrate is measured on lateral bipolar transistors, which show a bump in the beta curve, by using two different measurement methods. The first method makes use of the basewidth modulation technique [5,6]. The results are compared with lifetime values obtained from spectral response measurements. Both results are in close agreement.

The basewidth modulation technique is used on a parasitic PNP, formed by a p+ junction as emittor, the substrate as base region and the pwell as collector. To check the consistency of the method, measurements are repeated on devices with two different lateral basewidths, resulting in a current gain of respectively unity and two. The method consists of superimposing an AC voltage on the base-collector reverse voltage, producing a change in the width of the quasi neutral base region. An AC current can be measured in the collector due to a change of the slope of the minority carrier distribution in the base. The basewidth

modulation also gives a change in the substrate(=base) recombination current. The ratio of these AC currents, which is the ratio of the output- and transconductance (Go, Gr), multiplied by the transittime (τ_b), gives the lifetime. For an uniform doped base τb can be approximated by $W_b^2/(2\ D_p)$.

$$\tau = \tau_b \ Go / Gr$$

The method gives consistent values only if the measured Go and Gr are the result of the basevidening effect. To check this, Go and Gr are measured for different applied base emitter voltages. According to theory both curves have to be exponential with the Vbe.
Especially for Gr, the AC component of the base recombination current is rather small, while Vbe has to be kept constant during measurement, so that a closed loop measurement configuration was needed. The experimental results are given in figure 5, which shows the base and collector current, the Go and Gr parameters for different base emittor bias voltages and the ratio between Go and Gr. The Go/Gr ratio of 7 results in a lifetime of about 15 nsec.

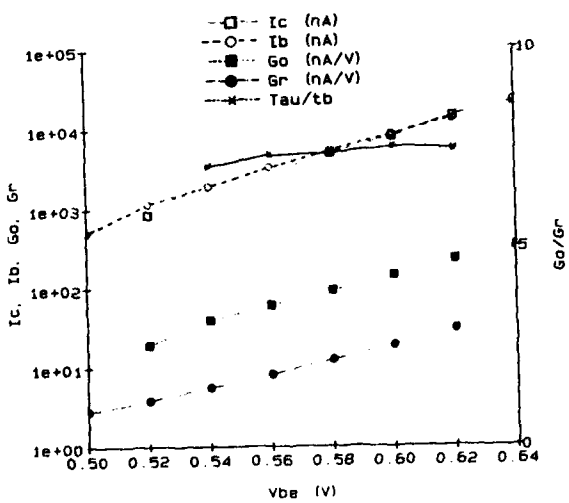


FIGURE 5: Collector-, base current, trans-, outputconductance and lifetime per unit base transit time for a pnp with 2μm basewidth.

Minority carrier lifetime can also be calculated out of the diffusion length obtained by spectral response measurements. For the long wavelength response of the photocurrent, following approximated relation exists between the diffusion length (L_p), the optical absorption coefficient (α) and the internal quantum efficiency (Q) :

$$\frac{1}{Q} = 1 + \frac{1}{\alpha\ Lp}$$

Extrapollation from the measured 1/Q versus 1/α curve, gives an intercept on the 1/α axis, which is equal to 1/Lp. The extrapollated value for Lp is about 5 μm which corresponds with a lifetime of 20 nsec.

Both methods lead to similar values for the lifetime in the base region of the parasitic PNP. The recombination lifetime is lower than 5 μsec, which results in a bump in the beta curve as measured.

REFERENCES

[1] R. D. Thornton et al., Characteristics and Limitations of Transistors, SEEC Vol. 4. Wiley, New York (1966)

[2] Sunlin Chou, Solid State Electronics, Vol 14, 811, 1971

[3] W. Shockley and W.T. Read, Phys. Review, Vol. 87, 5, 1952

[4] J. W. Slotboom and H. C. de Graaff, Solid State Electronics, Vol. 19, 857, 1976.

[5] A. Neugroshel, IEEE Trans. Electron. Devices, Vol 28, 108, 1981

[6] M.S. Birritella, A. Neugroshel and F.A. Lindholm, IEEE Trans. Electron. Devices, Vol 26, 1361, 1979

SHALLOW JUNCTION CONTACTS FOR LATCH UP RESISTANCE IN CMOS

F. Ruddell, E. Ling, B.M. Armstrong, H.S. Gamble and S.H. Raza

Department of Electrical and Electronic Engineering
The Queen's University of Belfast, Ashby Building
Belfast BT9 5AH. N. Ireland

An optimum $TiSi_2$ thickness of 55nM has been established for contacting 110nM P^+ layers. The P^+ layer sheet resistance is reduced to less than 3 ohms per square without compromising the reverse leakage currents of the junctions. The reduction in Gummel number for the P^+ layer has reduced the emitter efficiency of the junctions and yielded a factor of four increase in latch-up resistance.

1. INTRODUCTION

The scaling down of device dimensions for CMOS VLSI requires ultra shallow junctions with low sheet resistance. Whilst this is difficult to obtain, especially for the P^+ junctions of the P-channel transistors, several novel techniques have been proposed [1][2]. However the sheet resistance of P^+ layers, approximatley 100nM in depth, is of the order of 100-200 ohms per square. To reduce the sheet resistance it is common to form a self-aligned layer of titanium disilicide above the P^+ region.

During the formation of the titanium disilicide some of the boron doped silicon layer is consumed. There is therefore a limit to the maximum thickness that can be used. If the silicide layer is thick enough all of the P^+ layer is consumed and a $TiSi_2$-n-silicon Schottky barrier is formed. This results in a marked increase in both forward and reverse current for the junction and a reduction in the minority carrier current injected into the n-silicon substrate.

An additional requirement for Bulk CMOS is latch up resistance. Techniques for increasing latch up resistance focus on reducing the emitter-base shunting resistance or on lowering the gain product of the parasitic bipolar transistors. An alternative approach to reduce emitter efficiency is to systematically reduce the Gummel number of the $P^+$and N^+ layers forming the emitters of the parasitic bipolar transistors by formation of silicide contacts. There is therefore a need to establish the optimum thickness of silicide contacts to ultra shallow junctions to achieve low sheet resistance and enhanced latch up resistance, together with minimization of junction leakage, in order to take full advantage of a low power CMOS structure.

In this paper, diode reverse leakage currents and the minority carrier injection properties have been studied for 110nM P^+-N junctions with several titanium disilicide thicknesses. The ultra shallow junctions were produced by rapid thermal diffusion from a spin-on diffusion source [3]. Two processes have been designed. The first permits manufacture of low leakage P^+-N diodes for standard current-voltage characterisation. The second permits the manufacture of P-N-P bipolar transistors for emitter efficiency studies.

2. DIODE CHARACTERISTICS

P^+-N diodes were produced with a range of titanium disilicide thicknesses formed in the boron diffusion. The main features of the process are outlined in Figure 1. Mask design ensured that the Aluminium over the contact windows was smaller than the diffused area thus avoiding any field plate effects.

Typical values of reverse bias leakage

WAFER N-TYPE 5-10 ohm-cm <100>
BACK DAMAGE
1050°C 90 mins DRY OXIDE/TCE
REMOVE BACK OXIDE
1000°C 15 mins PHOSPHOROUS DIFFUSION
1000°C WET OXIDE (0.3 μm)
MASK 1 - DIFFUSION WINDOWS
CHEMICAL OXIDE (3 nm)
EMULSITONE BOROFILM 100 SPIN-ON DOPANT
PREBAKE 350°C 10 mins
1000°C 3 mins N_2 RAPID THERMAL DIFFN
REMOVE SPIN-ON GLASS
CVD OXIDE (0.3 μm)
MASK 2 - CONTACT WINDOWS

RF SPUTTER TITANIUM (76-80 μohm-cm)
620°C 6 secs $TiSi_2$ FORMATION
SELECTIVE ETCH
850°C 6 secs ANNEAL
500°C 10 mins FORMING-GAS ANNEAL
ALUMINIUM EVAPORATION
MASK 3 - PATTERN ALUMINIUM
ALUMINISE BACK

Figure 1

The P^+-N diode process flow

current density are shown in Figure 2 for silicide thickness ranging from zero to 60nM. It can be seen that there is little dependance of leakage current with silicide thickness for values less than approximately 40nM.

A large increase in leakage current was observed for a silicide thickness of 60nM.

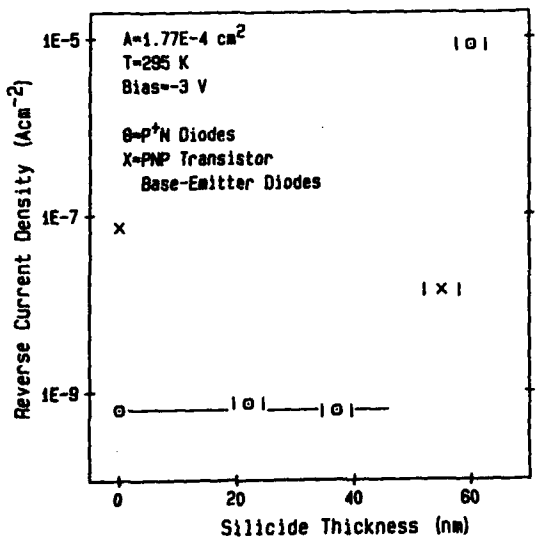


Figure 2

Reverse current density versus titanium disilicide thickness at a reverse bias of 3v for both P^+-N diodes and P^+-N emitter-base junctions.

Forward bias measurements revealed that these junctions were now operating as Schottky barrier diodes with a barrier height to n-silicon of 0.9v and an n factor of 1.24. This barrier height is considerably greater than that reported for $TiSi_2$-n-silicon and is due to the thin layer of boron remaining under the silicide which acts to enhance the built in potential barrier. The measured reverse leakage current is much higher than predicted by ideal Schottky barrier theory and may be attributed to localised roughness of the silicide.

3. MINORITY CARRIER INJECTION

Minority carrier injection ratios were measured using the P-N-P bipolar transistor structure shown in Figure 3. The n-base

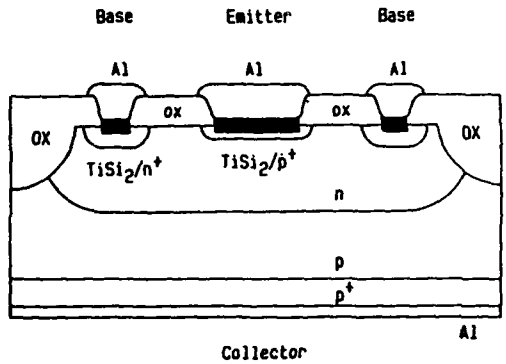


Figure 3

P-N-P transistor structure.

region was obtained by furnace diffusion from a spin-on phosphorus source. A reference standard transistor fabricated with a 0.6 micron thick emitter gave a transistor common emitter current gain of 200. All other transistors employed the 110nM boron junction and a range of silicide thicknesses from zero to 120nM were formed on this junction. An outline of the main process features is shown in Figure 4.

For zero thickness of silicide the transistor current gain has fallen to approximately 100. This can be attributed to

WAFER P-TYPE 5-10 ohm-cm <100>
1050°C 30 mins BASE DIFFN (PHOS SPIN-ON)
CVD NITRIDE
MASK 1 - PATTERN NITRIDE (DEFINE BASE)
PLASMA-ETCH SUBSTRATE
1150°C 105 mins WET OXIDE (LOCOS)
REMOVE NITRIDE
1150°C 20 mins WET OXIDE
MASK 2 - BASE CONTACT DIFFUSION WINDOWS
1000°C 15 mins PHOSPHOROUS DIFFUSION
REMOVE MASKING OXIDE
1150°C 20 mins WET OXIDE
REMOVE BACK OXIDE
1000°C 20 mins BORON DIFFUSION
MASK 3 - EMITTER DIFFUSION WINDOWS
FILMTRONICS BORON 'A' SPIN-ON DOPANT
1000°C 3 mins N_2 RAPID THERMAL DIFFN
REMOVE SPIN-ON GLASS
CVD OXIDE
700°C 10 mins N_2 DENSIFY OXIDE
MASK 4 - CONTACT WINDOWS

FINISH WITH $TiSi_2$/Al CONTACT STRUCTURE
AS FOR LOW-LEAKAGE DIODES

Figure 4

The P-N-P transistor process flow.

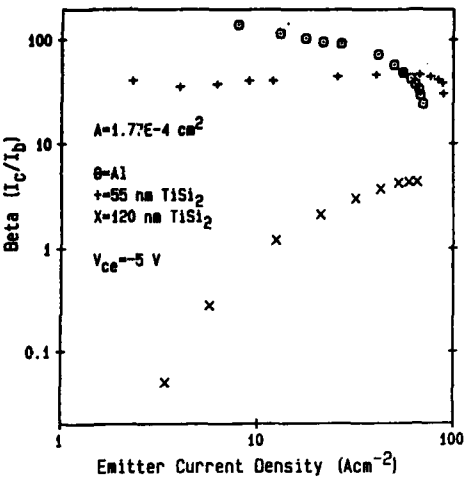


Figure 5

Transistor common emitter current gain versus emitter current density for a 110nM wide emitter with zero, 55nM and 120nM thicknesses of TiSi contact.

the fact that the emitter width of 110nM is now considerably less than the electron diffusion length in the emitter. Thus the reduction in emitter width has resulted in a 50% reduction in emitter efficiency compared to the reference long emitter transistor. The use of a 55nM silicide contact resulted in a current gain in the range 40-50. Again this is due to the further reduction in effectiye emitter Gummell number due to consumption of the silicon during silicide formation. The emitter efficiency has therefore been reduced to about 20-25% that of the reference transistor. A further increase in silicide thickness to 125nM ensured complete consumption of the boron diffused emitter. Current gains were reduced to 0.1 - ' for emitter current densities up to 10A.cm^{-2}. This large decrease in gain was due to the fact that the emitter-base junction now consisted of a Schottky barrier diode. Thus emitter efficiency has been reduced to less than 0.05% that of the reference transistor. The measured values of common emitter gain for each of these thicknesses of silicide is plotted as a function of emitter current density in Figure 5.

The forward and reverse biased characteristics of these emitter-base junctions were measured. The devices with zero and 55nM silicide thicknesses exhibited similar forward I-V characteristics and reverse leakage currents in the range 10^{-7} - 10^{-8}A.cm^{-2} at a reverse bias of -3V. These values, plotted in Figure 2 are in excess of those measured on the P^+-N diode test structures. However it should be noted that the P^+-N diodes were formed in relatively lightly doped substrates, and the process flow was specifically designed to reduce the influence of surface states, metallic impurities, lattice imperfections and phosphorous snowplow. The transistor test structures had base doping densities of the order 10^{17}cm^{-3} and the process flow was not optimised with respect to these effects. It is therefore important to note that with a 55nM silicide thickness the emitter-base operates as a P^+-N junction whose properties are not compromised by the formation of the silicide. This compares with the 60nM silicide formed on the P^+-N diodes where a

Schottky barrier was formed and the reverse leakage current was severely degraded.

The I-V characteristic under forward bias conditions in the emitter base junction with a 120nM contact is shown in Figure 6. The

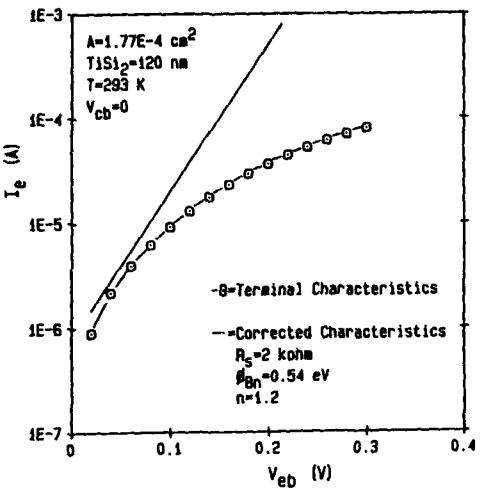


Figure 6

Forward bias characteristics of the emitter-base junctions contacted with 120nM of $TiSi_2$.

current levels measured were relatively high. The device terminal characteristics were therefore corrected to account for approximately 2kΩ of base resistance. The emitter-base junction was therefore characterised as a Schottky barrier diode with a barrier height to n-silicon of 0.54v and an n factor of 1.2. The barrier height is lower than reported for $TiSi_2$ - n-silicon and may be due to the fact that the silicide was formed by reaction with silicon heavily doped with boron. The contact therefore consists of both titanium disilicide and titanium boride.

The reverse bias leakage currents of these junctions exhibited values of the order 10^{-5}A. This is considerably higher than predicted by ideal Schottky barrier theory and can be attributed to the increased roughness of the thicker silicide and edge effects.

4. CONCLUSIONS

It has been established that the reverse bias leakage current properties of an ultra-shallow P^+-N junction are not compromised by the formation of a silicide contact until the junction becomes a Schottky diode. For the 110nM junctions employed in this work, the onset of Schottky operation occurs when a silicide thickness of between 55nM and 60nM is employed. The current densities of less than 10^{-9}A.cm^{-2} achieved on the diode test structures are much lower than other values reported in the literature for similar junction depths. The sheet resistance of these junctions can therefore be reduced to less than 3ohms per square using 55nM of titanium disilicide without compromising reverse leakage properties. An equally significant feature for CMOS circuits is that the emitter efficiency of the parasitic bipolar transistors can be reduced by decreasing the emitter Gummell number. The 110nM junction contacted with 55nM TiSi exhibited a 75% reduction in emitter efficiency and therefore a factor of four increase in latch up resistance as compared to a reference long emitter transistor.

ACKNOWLEDGEMENTS

F. Ruddell wishes to acknowledge the financial support of the Department of Education for Northern Ireland and STC Technology, Harlow, England.

E. Ling wishes to acknowledge the financial support of the Department of Education for Northern Ireland.

REFERENCES

[1] Wilson, R.G. J. Appl. Phys. (1983) 6879.

[2] Solmi, S., Landi, E. and Negrini, P., IEEE Trans Electron. Dev. Lett. (1984) 359.

[3] Ling. E., Maguire, P.D., Gamble, H.S. and Armstrong, B.M., IEEE Trans. Electron. Dev. Lett. (1987) 96.

THREE DIMENSIONAL DISTRIBUTION OF LATCH-UP CURRENT IN SCALED CMOS STRUCTURES

B3.1.3

Luca SELMI, Franco VENTURI, Enrico SANGIORGI, and Bruno RICCÒ

Department of Electronics, University of Bologna
Viale Risorgimento 2, 40136 Bologna, Italy

This paper presents a novel hysteresis phenomenon induced in the latch-up I-V characteristic of CMOS structures by three dimensional effects producing strongly bias dependent non-uniformities in the current lateral distribution. This behavior has been experimentally reproduced in a lumped element circuit, and a suitable model is presented.

1. INTRODUCTION

Latch-up represents a crucial problem to be solved in the scaling down of device dimensions toward the submicron range. In the study of this phenomenon it has been shown [1,2] that three dimensional effects play an important role in determining the latch-up behavior of common devices (usually described by the I-V coordinates of the holding and triggering points in the SCR characteristic), as well as their dependence on geometrical dimensions.

In this paper we will show that such effects can also produce unexpected but reproducible and well characterized hysteresis phenomena in common CMOS structures.

2. EXPERIMENTS

The devices used in the experiments were fabricated with a CMOS technology based on bulk p-substrate of 8×10^{14} cm^{-3} with 5 μm deep n-wells. Conventional four stripes test structures featured different top view aspect ratios (L=80,140 μm; W=50,100 and 200 μm) and p^+ to n^+ spacings (S=4,8 μm). The contacts on the p^+ emitters have been splitted in several equal parts (3 in the 50 μm, 5 in the 100 and 200 μm devices). Control devices without such a splitting were also fabricated and used to make sure that the results are not an artefact of the contact geometry and/or measuring circuit (fig.1).

During the measurements well and substrate terminals were shorted with the p^+ and n^+ emitters respectively, backplane substrate contact was kept floating, and the device was driven with a current source.

Fig.2 is a photograph of the I-V characteristic of a 200 μm structure (S=8 μm) without multicontact p^+ emitter. The hysteresis clearly visible around the holding point was found (below the holding point) also in devices with W=100 μm but not in the 50 μm structures.

The width of the hysteresis cycle, taking place at very low current and voltage values, is about 20 mV. This implies extremely small power level differences between stable limits of the hysteresis and allows to rule out the hypothesis that thermal or second breakdown effects are the causes of this behavior.

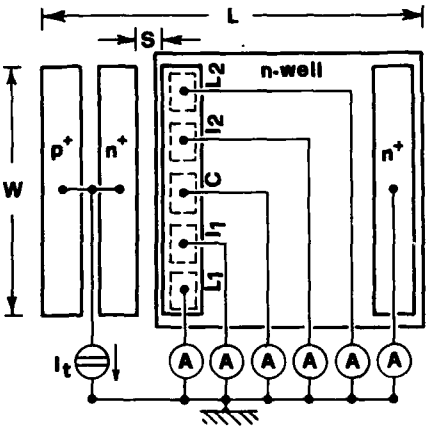


Fig.1 : Schematic representation of the multi-contact test structure layout and measuring circuit.

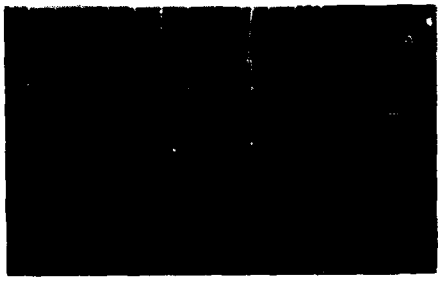


Fig.2 : I-V characteristic of a 200 μm structure without multicontact p^+ emitter (S=8 μm).

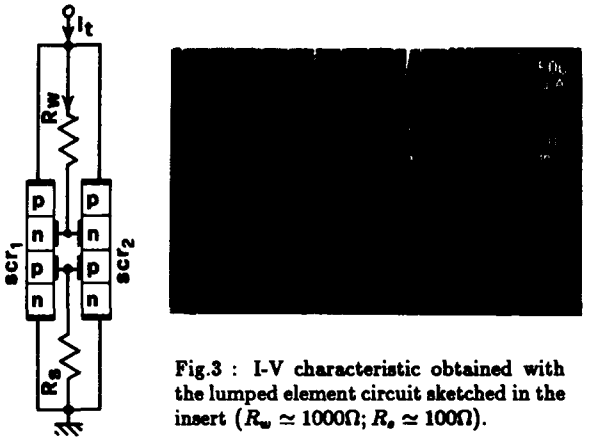


Fig.3 : I-V characteristic obtained with the lumped element circuit sketched in the insert ($R_w \simeq 1000\Omega$; $R_s \simeq 100\Omega$).

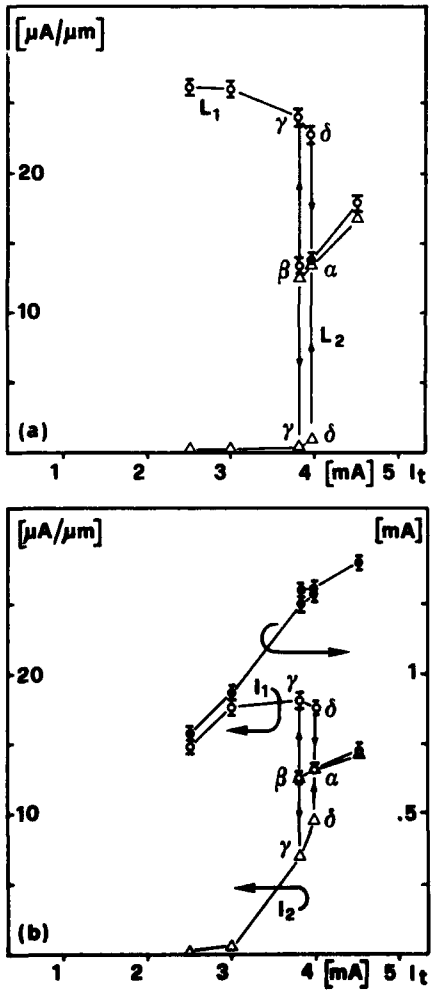


Fig.4a,b : Current density (per μm of W) through: a) the lateral p^+ emitter contacts (Δ,$\circ$); b) the intermediate p^+ emitter contacts (Δ,$\circ$) and the well contact ($\bullet$) vs. structure total current (S=8 μm,W=200 μm).

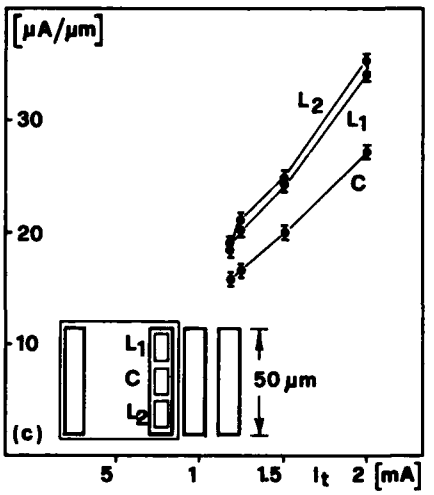


Fig.4c : Current density (per μm of W) through the p^+ emitter contacts of 50 μm wide structures (S=8 μm) vs. structure total current.

Fig.3, instead, shows a similar characteristic obtained with a lumped element circuit, sketched in the insert, based on two SCR's with a slight difference in the holding voltages ($\simeq$5 mV) connected in parallel . The presence of the hysteresis in the discrete circuit strongly indicates that two different devices are essential for the phenomenon to take place.

Fig.4 presents the current densities measured at various p^+ emitter sections of the 200 and 50 μm wide devices as a function of the total current (I_t) flowing across the structure. In particular curves in fig.4a,b refer to the lateral and intermediate emitters of the 200 μm device respectively, while the 50 μm stucture is reported in fig.4c. As can be seen the current excess at the structures edges at the holding point [1] becomes much larger with the device biased well in the conducting state. The difference between curves L_1 and L_2 as well as I_1 and I_2, increasing with total current, is probably a consequence of small parasitic contact resistance disuniformities.

The hysteresis instead is associated with completly different current distributions along the device width. In fact once the structure has reached the β point (where current splits almost simmetrically between lateral contacts) a further decrease in I_t causes current crowding at one side of the structure while the other rapidly drops out of latch-up (γ point). The comparison between fig.4a and b points out a current crowding enhancement below the holding point. The original simmetric situation can be obtained only increasing I_t above a threshold I_δ, higher than I_γ. Conversely

the 50 μm device (fig.4c) exhibits the same current distribution above and below the holding point.

3. DISCUSSION

The experiments, as the lumped element circuit indicates, can be interpreted by means of the structure shown in fig.3 composed of two SCR's (SCR_1 and SCR_2 respectively) connected in parallel and exhibiting slightly different I-V characteristics as a consequence of many possible contributions such as in particular distributed contact resistance. In this model for any given I_t the SCR's bias points $(I_1, V; I_2, V)$ must be such that $I_1 + I_2 = I_t$ while the presence of R_s and R_w provides a coupling term between the constituent SCR's.

Let us now discuss the structure behavior along the descending branch of the hysteresis cycle ($E \rightarrow \alpha \rightarrow \beta \rightarrow \gamma$ in fig.5b or $E \rightarrow H \rightarrow \alpha \rightarrow \beta \rightarrow \gamma$ in fig.5c), starting from a situation where the current distribution is essentially symmetric (as seen in the experiments).

With the SCR's in the low impedance condition a decrease in I_t causes a correspondent decrease in V until SCR_2 reaches its holding (Fig. 5a, point H_2) while SCR_1 is still in a positive differential resistance region (point A_1). If I_t is now further lowered there are two possibilities: a) a voltage $V > V_{H_2}$, exists such that $I_1(V) + I_2(V) = I_t$ and the SCR's are still in the latched state; b) the previous condition cannot be verified.

In the latter case (Fig. 5a,b) SCR_2 drops out of the latched state switching from H_2 to S_2 while the voltage across SCR_1 must be such that $I_1(V) = I_t$ (point C_1). Under these conditions the I-V curve of the complete structure (Fig. 5b) will exhibit a discontinuity ($\beta \rightarrow \gamma$ Fig. 5b) around its holding point as actually found in the 200 μm devices and experimentally reproduced with the lumped element circuit of fig.3.

If, instead, a voltage $V > V_{H_2}$ can be found so that $I_1(V) + I_2(V) = I_t$, SCR_2 will enter its negative resistance region ($H_2 \rightarrow R_2$) while the SCR_1 operating point moves from A_1 to B_1. This condition holds until a current level I_t^* is reached for which $g_1(I_1) + g_2(I_2) = 0$ and $I_1 + I_2 = I_t^*$ where g_1 is the differential conductance of SCR_1 (approximately constant and positive between points A_1 and B_1) and g_2 that of SCR_2 (negative but increasing between points H_2 and R_2). For $I_t < I_t^*$ SCR_2 drops out of latched state with effects similar to those of point b). While I_t decreases then, V reaches a minimum (point H in fig.5c), then increases slightly and finally switches on the other branch of the hysteresis cycle (γ).

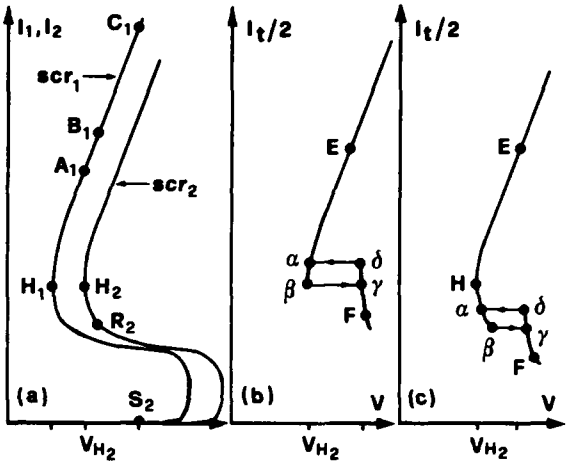


Fig.5 : Schematic representation of the I-V characteristics of: a) SCR_1 and SCR_2 (see insert of fig.3); b,c) the whole lumped element equivalent structure.

This model is confirmed by the lumped element circuit that furthermore allows to separately measure voltages and currents at every node thus exactly clarifing the way SCR_2 switches off. The results of the measurements on the actual distributed structures are shown in fig.4a,b and clearly indicate that when SCR_2 exits the latched state tens or hundreds of μA still flow through the off side of the p^+ emitter thus pointing out that the SCR_2 vertical transistor, that has higher shunting resistance, remains in its high gain, normal mode of operation but its collector current is not enough to keep the lateral npn in conduction. Finally the experiments show that currents flowing through the well and substrate contacts, hence the total emitter currents, suffer only minor changes during the $\beta \rightarrow \gamma$ transition.

Turning now to the other branch of the hysteresis cycle, i.e. that through the points $F \rightarrow \gamma \rightarrow \delta$, we deal with a circuit that operates very asymmetrically so that it is no longer possible to neglect the coupling of the two halves of fig.6 (or of the insert in fig.3), as implicetely done in the previous discussion. In the starting point F we have: T_2 off, T_1 on at low current level, T_3 and T_4 - representing the lower holding part of the device - heavily saturated. The lateral voltage drop due to majority carrier flow underneath the emitter diffusions, responsible of the different status between the left and right sides of the structure, is accounted in the model via $r_{s_2}, r_{s_4}, r_{w_1}$, and r_{w_2}. In particular at γ, r_{w_1} is crossed by the small T_1 base current while the others

by part of their transistors bypass currents. This situation is quite different from that of point β where currents still cross the structure simmetrically and T_2 base emitter voltage V_{be2} ($\simeq V_{be4}$), is the sum of two terms: the drop on R_s due to substrate shunting current and that on r_s due to the high T_1 collector current. In terms of voltages the situation at γ differs from that at β because the voltages across r_{w_1} and r_{s_2} are much smaller (while those on R_w and R_s have not changed substantially, as the overwhelming part of the shunting currens now flows through the heavily saturated T_3 and T_4). In particular V_{be2} is now much smaller than V_{be4}, even if T_1 is not really switched off, since its collector current is in any case much smaller than that of T_3.

If the total current is now increased, the starting value of V_{be2} is too small for T_2 to turn on and a positive feedback with T_1 to be established. The threshold for this to be the case will be reached again only when I_t reaches a large enough value I_δ. Consequently between γ and δ the structure I-V characteristic is substantially different from that of the branch $\alpha \rightarrow \beta$.

Of course the actual resistances values can deeply influence the device behavior. In particular r_{s_i} (r_{w_j}) should be sufficiently high compared to R_s (R_w) to induce an appreciable difference in the status of T_2 and T_4, as necessary for the hysteresis to occur. This has also been verified experimentally as an addition of external resistances increasing the R/r ratio makes the hysteresis disappear in all cases. Insufficient values of r_{s_i} and r_{w_j}, due to its reduced width, explain, in our opinion, why no hysteresis can be found in the 50 μm devices.

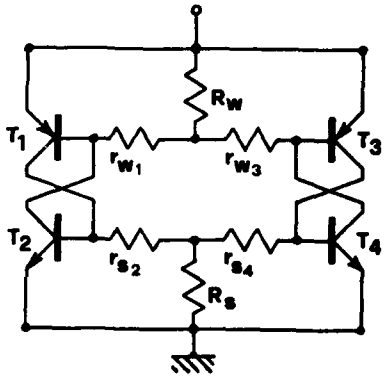


Fig.6 : Lumped element equivalent circuit of the 200 and 100 μm structures.

4. CONCLUSIONS

A novel hysteresis phenomenon caused by three dimensional effects and producing strongly bias dependent nonuniformities in the current lateral distribution was observed in the latch-up I-V characteristic of CMOS structures. This behavior has been experimentally reproduced in a lumped element circuit and characterised by a suitable model. Furthermore it is shown that in wide structures (sufficiently large to be sketched as composed of different sections connected in parallel) bistability phenomena can take place (due to slight disuniformities and/or enough large trasversal resistive coupling between the various sections).

ACKNOWLEDGEMENTS

The authors would like to thank Giuseppe Corda and Alfonso Maurelli of SGS-Microelectronics (Milan) for their technical cooperation and support in realising and processing the devices used in this work.

REFERENCES

[1] E. Sangiorgi, B. Riccò, and L. Selmi, "Three dimensional distribution of CMOS latch-up current", IEEE Electron Device Lett., vol. EDL-8, p.154, 1987.

[2] A. G Lewis, R. A. Martin, T. Y. Huang, J. Y. Chen, "Three-dimensional effects in CMOS latch-up", IEDM Tech. Dig., p.248, 1986.

A NEW SCR PARAMETER EXTRACTION METHOD TO HELP DESIGN FOR RELIABILITY IN CMOS CIRCUITS

K.ERDELYI
Research Institute for Technical Physics of the Hungarian Academy of Sciences
H-1325 Budapest, P.O.Box 76. Hungary

G.KNAPP
Microelectronics Company
H-1325 Budapest, P.O.Box 21. Hungary

1.INTRODUCTION

The parasitic SCR action in bulk CMOS integrated circuits has been extensively studied. Although there are several experimental methods to determine and study the latch-up sensitive parts of a layout, the prediction of latch-up behaviour is strongly desired.

CAD oriented techniques were developed to extract latch-up sensitive parts of a CMOS layout (1) thus reducing the problem to electrical simulation. However optimization of the model complexity is required because of the large number of possible paths and the effort to reduce computational time and cost (2). The prefered lumped element equivalent circuits have the disadvantage of inaccuracy in comparison with the 2D and 3D models since the parasitic structure can hardly be approximated by a network of individual elements (3). In spite of this there is still a need for one-dimensional SCR model which can be implemented into a general purpose circuit simulation program.

In the following sections we present our results on SCR parameter extraction by fitting the Gummel-Poon model on discrete bipolar transistor characteristics and a modified parameter extraction method that offers a better description of the SCR on-state characteristic.

2.GUMMEL-POON MODELING

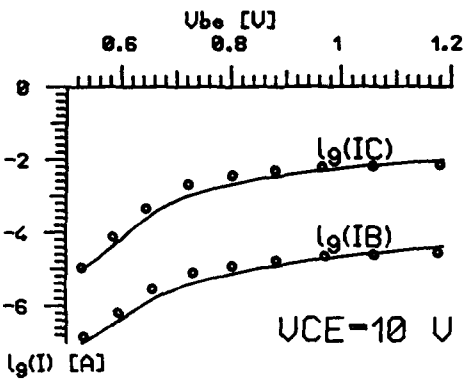
Parameter extraction begins with determination of initial values and is completed by a least-square fit on input, transfer and output characteristics. To avoid irreal final parameter values or divergence,the initial values should be carefully selected. Parameter determination technique had to be slightly modified as the effects of unusual parameter values of parasitic bipolars make difficult to separate the different regions of operation. The $\lg(I_C/I_B)$ versus $\lg(I_C)$ in normal active mode (common emitter configuration) proved to be efficient to determine the termination of generation-recombination dominant region and the beginning of the high level injection region as the collector current increases. From Gummel-Poon equations a formula for I_C/I_B can be derived. For simplicity the Early effect is not included (eq.1.).

$$\text{eq.1}\quad \frac{I_C}{I_B}=\frac{B_F}{\left[1+\frac{I_{SE}}{I_S}B_F\left(\frac{I_C}{I_S}\right)^a\right]\left[1+\frac{I_C}{I_{KF}}\right]};\quad a=\frac{N_F}{N_E}-1$$

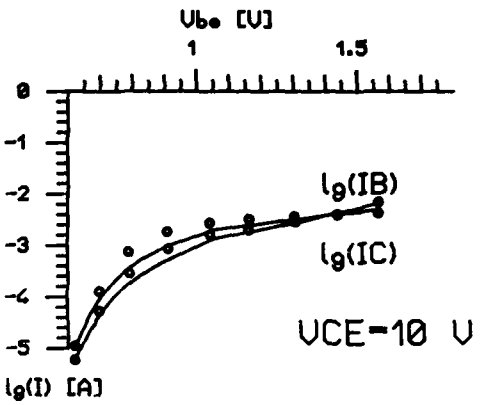
where BF maximum forward beta
ISE B-E leakage saturation current
IS transport saturation current
NE B-E leakage emission coefficient
IKF corner for forward beta
high current roll-off
NF forward current emission coefficient

Note that the equation is independent of resistances. Applying logarithmic scale a slope of (1-NE/NF) can be derived for the former and -1 for the latter region.

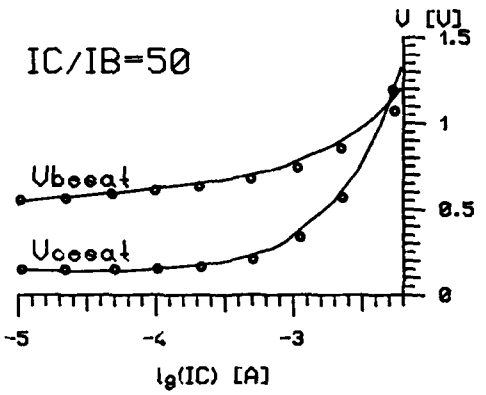
Transfer characteristics were used to determine parameters responsible for normal active behaviour (eq.2,3)(SCR off-state) and saturation curves to characterize the device with



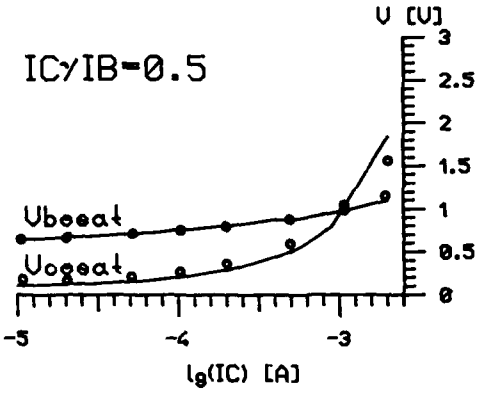
1.a.



1.c.



1.b.



1.d.

Fig.1. Bipolar transistor characteristics
1.a. npn transfer and input characteristics
1.b. npn C-E, B-E saturation voltage vs. collector current
1.c. pnp transfer and input characteristics
1.d. pnp C-E, B-E saturation voltage vs. collector current

eq.1 $$\lg I_C = \lg I_S + (V_{BE} - I_B R_B - (I_B + I_C) R_E)/(N_F V_T)$$

eq.2 $$\lg I_B = \lg \frac{I_C}{B_F} + \lg\left(1 + \frac{I_{SE} B_F}{I_S} \exp \frac{(N_F - N_E)(V_{BE} - I_B R_B - (I_B + I_C) R_E)}{N_F N_E V_T}\right)$$

eg.3 $$V_{CESAT} = V_{CES0} + I_C \left(R_C + \frac{N_R (B_S + 1)}{N_F B_S} R_E\right)$$

eq.4 $$V_{BESAT} = I_S \left(1 - \frac{B_R + 1}{B_R} \exp\left(-\frac{V_{CES0}}{N_R V_T}\right)\right) \exp \frac{V_{BE} - (R_B + (B_S + 1) R_E) I_C / B_S}{N_F V_T}$$

where
N_R reverse emission coef.
B_R maximum revese beta
$B_S = I_C / I_B$

both junctions forward biased (eq.4,5) (SCR on state).

A modified gradient method was used for efficient optimization of parameters. The comparison of calculated and measured characteristics are shown in Fig.1.

The calculated parameters (capital letters) can be expressed as follows (see Fig.3.):

$R_{CP}=r_W+r_{CP}$ $\quad R_{CN}=r_S+r_{CN}$

$R_{BP}=r_S+r_{BP}$ $\quad R_{BN}=r_W+r_{BN}$

$R_{EP}=r_{EP}$ $\quad R_{EN}=r_{EN}$

For first order SCR modeling input parameters were determined with the assumption of positive resistances;

$r_W = \min(R_{CP}, R_{BN})$

$r_S = \min(R_{CN}, R_{BP})$

The other parameters were derived in compliance with the selection of r_W and r_S.

We performed SCR simulation with these extracted parameters. In spite of the fairly good parameter fitting of bipolars, it was found that the measured and calculated SCR data differ significantly especially in the on-state region Fig.2.

3.MODIFIED SCR PARAMETER EXTRACTION

In order to get better description of latch-up hardness of CMOS structure our main idea was to extract the model parameters from the measured on state characteristics. The parameter V_{min} (the minimum voltage required to sustain on-state) is prefered to the commonly used I_{hold} and V_{hold} since the breakpoint occurs as the sum of the generator impedance and the negative impedance of the structure becomes zero, so these latter parameters are not free from measurement scheme dependence.

To obtain data in the on-state operation region I_W-I_{tot}, I_S-I_{tot} and V_{tot}-I_{tot} characteristics were measured (Fig.3.) according to (4). The selection of these parameters seemed to be useful to describe the coupling effect between bipolar structures.

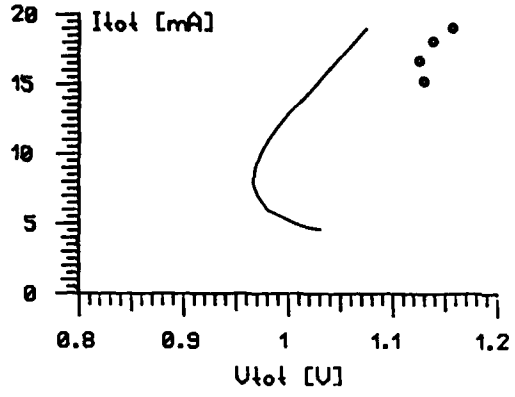


Fig.2. SCR I-V characteristic
measured: dashed line
simulated: o

A non-destructive DC measurement technique was developed for Keithley measuring system. Current generator was used as a supply. A current value greater then the holding current was set in stand-by state. Latch-up was triggered by displacement current initiated by the voltage spike at turning the generator on. The forced current was decreased until the the compliance limit of the generator was sensed at the device terminals. I_W, I_S and V_{tot} were measured during the process.

The measurements were carried out on special test structures containing the latch-up sensitive part of a CMOS inverter with different geometries.

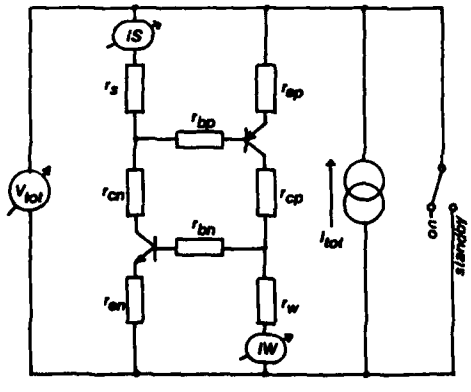


Fig.3. Measurement set-up.

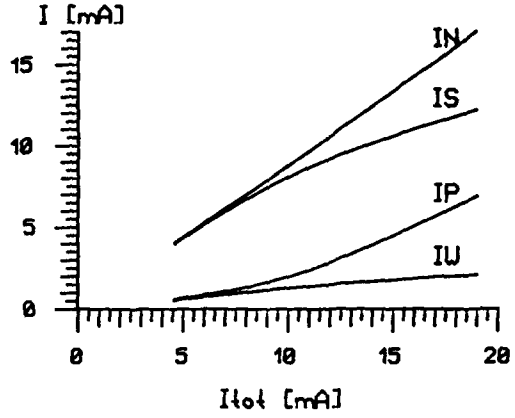


Fig.4. Measured terminal currents

On the basis of these data the same fitting procedure was performed as in the case of bipolar parameter extraction with difference in the error function calculation:

$$ERR = (I_{CN}+I_{CP}-I_{tot})^2/I_{tot}^2$$

I_{CN} and I_{CP} are the collector currents of the n- and p-type bipolar transistors respectively. These currents were determined by Gummel - Poon equations.

$$I_C = I_C(V_{BE}, V_{BC}, \underline{x})$$

where $V_{BE} = I_S R_S$ for p-type and $I_W R_W$ for n-type transistors

$V_{BC} = I_S R_S + I_W R_W - V_{tot}$ for both transistors

$\underline{x}$ repersents the parameters of transistors.

Since I_C is an implicit function an internal iteration was applied. The results of the bipolar parameter extraction were used as initial parameters for the fitting procedure. The calculated and measured SCR values are compared in Fig.5.

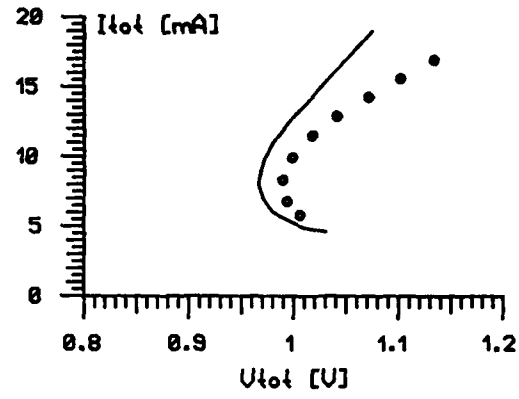


Fig.5.SCR I-V characteristics
measured: solid line
simulated: o

4.CONCLUSION

The parameter of SCR model based on fitting on on-state characteristics provides better agreement with the measured data.However the parameters obtained differ significantly from those determined from measurements of bipolar transistors.The geometry dependence of the extracted parameters needs further investigation.

ACKNOWLEDGEMENT

The authors would like to thank M.Seres his help in applying the modified gradient method for parameter fitting.

REFERENCES

(1) Faricelli, J.,Frey, J.
Symp.on VLSI Technology 1982
pp.78-79

(2) Wieder, A.,Harter, J.,Werer, C.
IEEE Trans.on El.Dev.Vol.ED-30.
pp.240-245. 1983.

(3) Ochoa, A.Jr,Dessendorfer, P.V.
IEEE Trans.on Nucl.Sci.Vol.NS-28.
pp.4294-4294. 1981.

(4) Hu, G.J.
IEEE trans.on El.Dev.Vol.ED-31.
pp.62-67.1984.

WEDNESDAY, SEPTEMBER 16, 1987

SALA BIANCA - CHAIRMAN: H. WIEDER

9.45 - 10.45
11.45 - 12.30 Panel on «Very High Speed Integrated Circuits» (Availability, Perspectives, Applications)

Organized by: Commission of the European Communities DG XIII - ESSDERC Committee

Members:

D. Baker	BTRL (GaAs Systems)
J. Borel	Thomson/SGS (Si Technology, Systems)
T. Ning	IBM (Si Technology)
G. Pelous	CEC (Technology)
M. Rocchi	LEP (GaAs Technology)
H. Stegmeier	Siemens (Systems)
B. Wilson	Plessey (Technology, Systems)
A. Wieder	Siemens (*Chair*, Si Technology, Systems)

Procedure:

The discussion will be initiated by short statements from the panel members.

Topics:

Technologies
(Availability, Perspectives, Applications)

Applications
(Status and Trends; Impact of Technology on Products and vice versa)

Session D3.1

Laser
I

Chairman: C. Wood

Wednesday, September 16, 1987

DEGRADATION BEHAVIORS IN LPE GROWN DFB LASER DIODES

Mitsuo FUKUDA, Masamitsu SUZUKI, George MOTOSUGI, and Tetsuhiko IKEGAMI

NTT Electrical Communications Laboratories, 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa, 243-01 Japan

Jun-ichi YOSHIDA

NTT Research and Development Bureau, 3-9-11, Midori-cho, Musashino-shi, Tokyo, 180 Japan

Degradation modes of 1.3 μm and 1.55 μm buried heterostructure (BH) distributed feedback (DFB) lasers are discussed and several degradation modes are clarified. DFB lasers essentially fail due to BH interface degradation in the same manner as BH Fabry-perot (FP) lasers. Additional degradation modes peculiar to DFB lasers are found to exist.

1. INTRODUCTION

DFB lasers have been found to be promising single wavelength light sources for various optical fiber transmission systems. Reliability is a very important factor in the application of such lasers to transmission systems. This paper discusses failure modes in DFB lasers, including a device lifetime comparison between 1.3 μm and 1.55 μm InGaAsP/InP DFB lasers.

2. SAMPLE PREPARATION

The 1.3 μm and 1.55 μm DFB lasers used in this study are BH types without a phase shifter in their cavity. The cavity length was about 300 μm and stripe width was 1.0-2.0 μm. These lasers were obtained from liquid phase epitaxal (LPE) wafers through the cleavage process. These wafers were fabricated as follows: (1) InGaAsP/InP double heterostructure (DH) wafers were etched down the mesa structure and (2) then the mesa side walls were buried with a p- and n- type InP layers, and an n-type InGaAsP layer by LPE. The typical threshold current of these DFB lasers is about 20-30 mA at room temperature.

3. AGING TEST AND RESULTS

3.1. Degradation Modes

In general, etched-mesa-buried-heterostructure FP lasers are degraded by BH interface degradation between first step growth layers (DH layers) and second step growth layers (Burying layers) as shown in Fig. 1 [1]. This interface degradation mainly depends on the injected current density rather than light output power and these degradation modes were

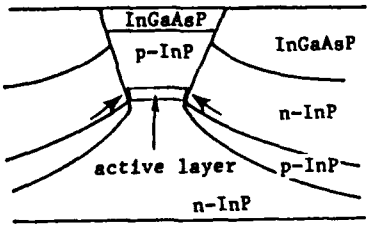


FIGURE 1

Degradation of BH lasers. Two arrows indicate the degraded areas.

clarified in detail [1]. Interface degradation can also be observed in DFB lasers. In addition, a degradation mode peculiar to DFB lasers exist which severely affect their aging characteristics. DFB lasers with super liner current-light output characterisctics shown in Fig. 2 tend to be unstable during aging. Lasing wavelength gradually gets shorter or longer depending on the degree of degradation. This phenomenon is mainly caused by inhomogeneous optical field distribution in the active region due to inhomogeneous injected current flow.

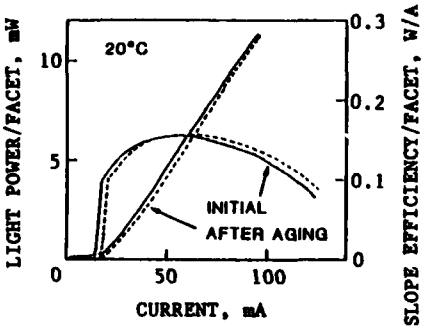


FIGURE 2

Change in lasing characteristics of 1.55 μm DFB laser. Aging for 20 h was carried out under a constant current of 200 mA.

These degradation modes become serious problems in terms of reliability, if the lasers which are severely degraded by above discussed modes are applied to transmission systems. However, those lasers can be screened by suitable pre-aging [2].

3.2. Long Term Aging

Aging tests were carried out on 1.3 μm lasers under a constant power of 8 mW/facet at 50°C (average driving current : 70 mA) [3]. The front facet of these lasers was coated with an anti-reflecting (AR) film. The results of the aging test are shown in Fig. 3. Before this aging test, the devices were subjected to a two-step screening test and only lasers which suffered little degradation were selected for the long term aging test. The average increasing rate of driving current was estimated to be 5×10^{-6} h^{-1}, and this value is small enough to assure long life of DFB lasers. These aging tests confirmed that there is no change in the lasing wavelength.

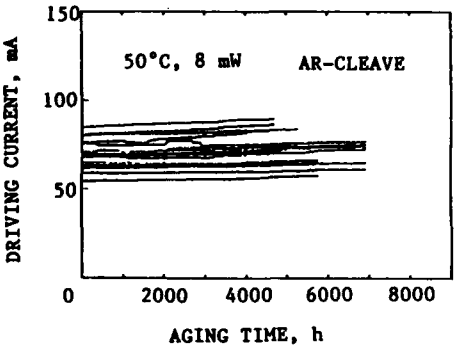


FIGURE 3

Aging characteristics of 1.3 μm DFB lasers.

For 1.55 μm DFB lasers with an AR-coated facet, the same screening tests and preliminaly aging tests were performed. The aging was carried out under a constant power of 5 mW/facet at 50°C (average driving current : 80 mA). The average increasing rate of driving current was estimated to be 1×10^{-6} h^{-1}. The wavelength of these lasers was also unchanged.

The increasing rate of driving current for 1.55 μm lasers was smaller than that for 1.3 μm lasers, even though the driving current for 1.55 μm lasers was larger than that for 1.3 μm lasers. This fact can be understood as an BH interface quality difference. The BH interface for 1.55 μm lasers is cleaned naturally by the melt-back process during second step growth and the melt back process hardly occurs for 1.3 μm lasers [4]. (Here, 1.3 μm lasers, where their

BH interfaces were artificially cleaned by m..c-back, were hardly degraded under aging [4].) This resulted in the lower increase in driving current for 1.55 μm. Under severe aging conditions(25°C, 20 mW/facet), it is also confirmed that BH interface degradation for

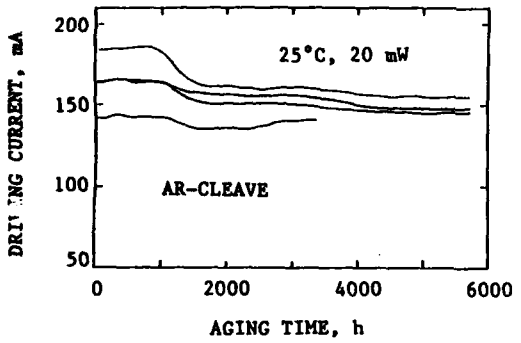


FIGURE 4

High power aging characteristics of 1.55 μm DFB lasers at 70°C.

1.55 μm lasers is small, as shown in Fig. 4. These results demonstrate the potential for high power operation of 1.55 μm DFB lasers without a large increasing in the driving current.

4. CONCLUSION

Several degradation modes in DFB lasers were clarified. DFB lasers essentially fail due to BH interface degradation in the same manner as FP lasers. In addition, degradation modes peculiar to DFB lasers were found to exist. Through suitable screening processes, DFB lasers can be selected for use in optical transmission systems.

ACKNOWLEDGEMENTS

The authors would like to thank K. Kurumada, N. Miyahara, and N. Tsuzuki for their valuable comments and encouragement throughout this work.

REFERENCES

[1] Fukuda, M. and Iwane, G., J. Appl. Phys. (1985) 2932.

[2] Ikegami, T., Takahei, K., Fukuda, M., and Kuroiwa, K., Electron. Lett. (1983) 282.

[3] Motosugi, G., Suzuki, M., and Saruwatari, T., Rev. of ECL, NTT (1987) 239.

[4] Fukuda, M., Noguchi, Y., Motosugi, G., Nakano, Y., Tsuzuki, N., and Fujita, O., IEEE J. Lightwave Tech., in print.

1300 nm DFB-LASER WITH REACTIVE ION BEAM ETCHED GRATING DEFORMATION-FREE OVERGROWN BY LPE

M. Schilling, K. Wünstel, H. Schweizer, J. Scherb, A. Mozer, K. Lösch, O. Hildebrand

Standard Elektrik Lorenz AG
Research Centre, Optoelectronic Components Division ZT/FZWO
Lorenzstr. 10, D-7000 Stuttgart 40, Federal Republic of Germany

DFB lasers were fabricated for the first time with gratings dry etched into InP substrate by reactive ion beam etching (RIBE). Deformation-free overgrowth of these 0.2 μm deep corrugations was performed by liquid phase epitaxy (LPE). Threshold currents (cw,25 °C) as low as 23 mA demonstrate that no severe degradations occur in these devices due to surface damage caused by the dry etching. Single mode operation with a sidemode suppression > 35 dB is obtained due to strong coupling of the grating. HF-measurements reveal a 3 dB modulation bandwidth of 9 GHz for the DFB-DC-DCPBH structure.

1. INTRODUCTION

Distributed feedback (DFB) lasers with stable single longitudinal mode operation are required for optical fibre transmission systems with high bit rates over long distances.
However, the fabrication of these devices involves very critical processing steps like diffraction grating formation and overgrowth of the fine corrugations.

Submicron gratings formed by wet etching typically show low depth to width aspect ratios. Their profiles are determined by preferential etching of special crystallographic planes. This leads to unfavorable grating shapes which are on the one hand often difficult to overgrow, and result on the other hand in weak grating coupling strength, especially when additional degradation occurs during the epitaxial process.
Recently we have shown that the application of reactive ion beam etching (RIBE) in connection with effective suppression of thermal deformation before liquid phase epitaxial (LPE) growth is a promising way to overcome the above mentioned problems and to realize the required submicron structures in the InP material system [1,2].

In the present work this technique is successfully applied to the fabrication of DFB-DC-DCPBH lasers in the 1.3 μm wavelength region.

2. DEVICE TECHNOLOGY

The second-order diffraction gratings with a period around 400 nm are formed in photoresist on top of (100) oriented n-doped InP substrates by holographic lithography. These patterns which are aligned along the [01$\bar{1}$]- direction are transferred into the InP wafer by a dry etching process. A mixture of 95% argon and 5% oxygen is used for the reactive ion beam etching (RIBE) performed in a commercial system with a Kaufman type ion source (TePla, Ribetch 160) as described [2]. Grating structures up to a depth of 0.25 μm are obtained in this way. Afterwards the resist mask is removed in an oxygen plasma followed by cleaning in organic solvents and a thorough oxide removal. No additional wet etching of the semiconductor surface or anodization step as reported in [3] is necessary to dissolve the residual damage layer.

In a first liquid phase epitaxial step the n-InGaAsP waveguide layer (d_w = 0.2 μm, λ = 1.15 μm) is grown directly on the corrugated substrate, followed by the undoped active InGaAsP layer (d_a = 0.12 μm, λ =

1.3 μm) and a p-InP confinement layer (d_c = 0.5 μm) completing the basic DH-laser wafer (Figure 1).



FIGURE 1

SEM cross-section of a laser wafer with dry etched grating after first LPE step.

The excellent embedding features of LPE allow for the full planarization of the deep corrugations by the waveguide layer and guarantee the planar growth of the active layer. Applying the GaAs-cover technique during low temperature LPE we obtain completely degradation-free overgrowth [2] of the argon/oxygen-RIBE etched trapezoidal gratings as demonstrated in Fig. 1. Finally, the DFB wafers with dry etched gratings are processed to index-guided DC-DCPBH lasers [4] suitable for high speed operation.

3. RESULTS

The light/current characteristic of a DFB-DC-DCPBH laser incorporating a dry etched deep grating is shown in Fig. 2. The threshold current I_{th} of this 180 μm long chip amounts to 29.5 mA under cw-operation at 25°C. A quantum efficiency of 18% (b.f.) is measured for this device. Threshold currents range from 23 mA to 31 mA for 140 μm to 220 μm long lasers cleaved from the same wafer. Quantum efficiencies up to 28% are obtained.

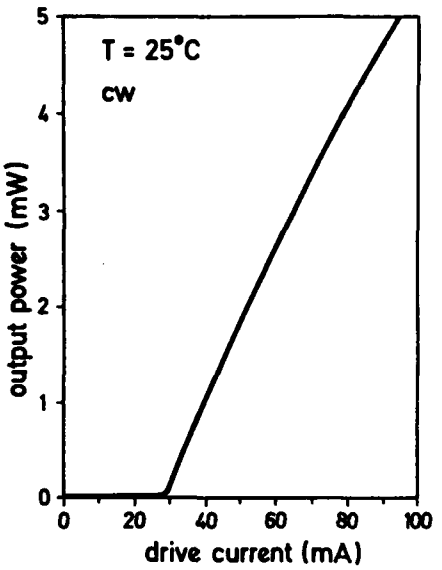


FIGURE 2

Light-current characteristic of a DFB-DCPBH laser fabricated from a wafer according to Fig.1.

Single mode operation with a strong side-mode suppression better than 35 dB is observed. An example of these measurements in the temperature range from 10°C to 40°C is shown in Fig. 3. In this figure the mode spectra (log scale) of an as-cleaved laser without antireflection coating are plotted for various temperatures at constant drive current I = 2.7 I_{th} (25°C). The measured emission wavelength shift with temperature is 0.09 nm/K which is typical for DFB operation corresponding to the temperature dependence of the refractive index.

Coupling coefficients κ of DFB-lasers are often determined from subthreshold stopband-width measurements. However, due to the dominating role of the DFB mode even below threshold a direct evaluation of κ is very difficult for the present lasers. From the depth and shape of the grating an estimation for κ based on Streifers model [5] gives coupling coefficients well above 100 1/cm.

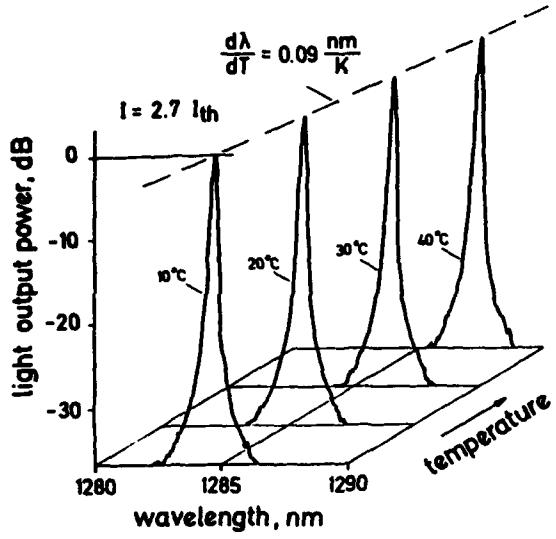


FIGURE 3

Mode spectra as a function of temperature at constant drive current.

To test the HF properties of our lasers, measurement of the frequency response under sinusoidal small-signal modulation was performed. The results are depicted in Fig. 4 as a function of different bias levels corresponding to various optical output powers. A maximum 3 dB modulation bandwidth of 9 GHz is measured at $I = 2.4\ I_{th}$ indicating that the laser is suitable for transmission at high bit rates.

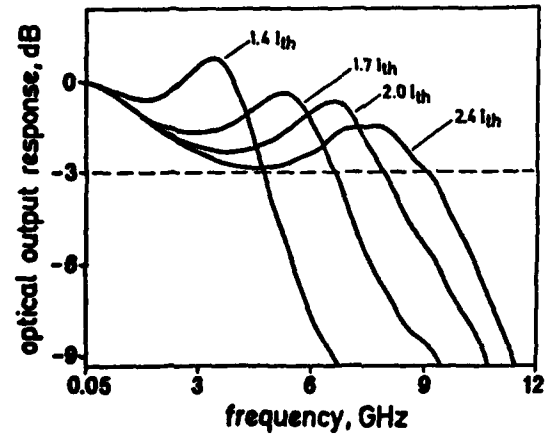


FIGURE 4

Small-signal frequency response for different bias levels.

4. CONCLUSIONS

DFB-Lasers were realized for the first time with gratings dry etched into InP-substrate as deep as 0.2 μm by reactive ion beam etching (RIBE). Deformation-free overgrowth of these deep corrugations followed by the complete laser structure was performed by liquid phase epitaxy (LPE). Threshold currents as low as 23 mA at room temperature under cw operation indicate that no severe degradations occur in these devices due to the surface damage possibly caused by the dry etching. From the low threshold values obtained we conclude that crystal damage as created by RIBE is in-situ annealed during the LPE process. Therefore we can fully profit by the advantages of the sharp edged and very deep dry etched gratings resulting in a strong coupling with a sidemode suppression better than 35 dB over a large temperature range.

In addition, excellent high frequency characteristics due to the DC-DCPBH structure with a 3 dB modulation bandwidth of 9 GHz are observed.

All the data presented above are first results obtained from devices which are not yet fully optimized. Nevertheless, the capability of this type of lasers for optical fiber communications is clearly demonstrated.

ACKNOWLEDGEMENTS

We wish to thank G. Müller, B. Knapp, A. Rudolf, G. Jahner and F. Schuler for support in epitaxy, device processing and characterization.

This work was funded by the German ministry of research and technology (BMFT).

REFERENCES

[1] Winstel, K. and Schilling, M., paper 63, 16th ESSDERC, 9.Sept. Cambridge, UK Europhysics Conf. Abstr. Vol. 10G (1986) 148

[2] Schilling, M. and Winstel, K., Appl. Phys. Lett. 49 (1986) 710

[3] Temkin, H., Dolan, G.J., Olsson, N.A., Henry, C.H., Logan, R.A., Kazarinov, R.F. and Johnson, L.F., Appl. Phys. Lett. 45 (1984) 1178

[4] Winstel, K., Mozer, A., Schilling, M., Luz, G., Lösch, K., Schweizer, H., Schesmel, G. and Hildebrand, O., Electron. Lett. 22 (1986) 1144

[5] Streifer, W., Scifres, D.R. and Burnham, R.D., IEEE J. Quantum Electron. QE-11 (1975) 867

1.5 µm DFB-BH LASER GROWN BY HYBRID LPE-MOVPE GROWTH

D. Lesterlin, J. Charil, B. Rose, M. Gilleron, P. Correc, J.C. Bouley

Centre National d'Etudes des Télécommunications
Laboratoire de Bagneux
196 avenue Henri Ravera - 92220 BAGNEUX - FRANCE

1.5 µm distributed feedback buried heterostructure lasers have been fabricated by a two step hybrid LPE-MOVPE growth. A threshold current of 38 mA and a single mode DFB operation up to 8 mW in CW have been obtained with a new stripe design.

1. INTRODUCTION

DFB Lasers are promising light sources for future optical communication systems. Their capability of maintaining longitudinal single mode operation under high frequency current modulation permits an increase of the bit-rate. Their narrow linewidth is also well suited for both coherent systems and wavelength multiplexing.

In order to improve the yield and the reproducibility of laser fabrication, MOVPE techniques tend to replace conventionnal LPE growth. MOVPE gives better uniformity in composition as well as in layers thicknesses than LPE[1-2].

The use of MOVPE has been previously reported for the realization of ridge DFB lasers[3-4]and BH-DFB lasers[5]. In the latter case, the fabrication involves three epitaxial steps over a corrugated substrate which implies the use of a first order grating to obtain low threshold current[6]. In our case, we propose a BH-DFB laser structure realized in two epitaxial steps with the use of a second order grating etched into the quaternary guide-layer.

2. MOVPE REGROWTH OVER GRATING

In the case of DFB-Lasers, MOVPE also minimizes the grating deformation during epitaxial regrowth[3], this is of great interest because the grating shape defines the feedback strength through the effective index variation it implies.

Two reasons could be invoked to explain that preservation :

First, the heat treatment before growth, where most of the deformations occur according to [7], is very short in an MOVPE technique compared with the homogeneisation period needed in the LPE process. This prevents any thermal deformation, providing some cares will be taken in the gas environment during heat.

Secondly, vapor phase epitaxy is an irreversible phenomena excluding any redissolution of the grating at the beginning of growth as it could take place in the quasi-equilibrium liquid phase epitaxy.

Moreover, in the structure we proposed, the grating etched in the quaternary layer above the active region, is less sensitive to thermal deformation than one etched in InP because the alloy contains less phosphorus , the most volatile compound among the constituants.

However, the gas ambient before MOVPE grating overgrowth is very critical and there is a controversy reported in the literature : [3] claims that PH_3 partial pressure enhances grating deformation before overgrowth when [6] uses a PH_3/AsH_3 mixture to suppress the thermal decomposition of the corrugated layer. So, we have performed some experiments to define the gas environment needed to protect the surface corrugation during our epitaxial process.

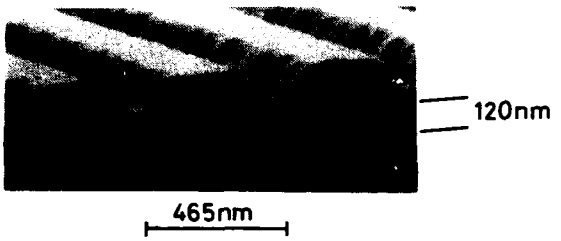


FIGURE 1
Etched grating in InGaAsP layer

Figure 1 shows a scanning electron microscope (SEM) photography of a grating made by classical holographic exposure and chemical etching (a Si_3N_4 layer covers it for lithographic purpose as explained in the third section). The period is 465 nm and the depth is 120 nm. We have done two separated InP regrowths by atmospheric pressure MOVPE on part of the wafer. Figures 2a and 2b show the grating profile after 1 µm InP growth. In case a, 1.2% PH_3 by volume was introduced in the carrier gas whereas, in case b, only 0.12% PH_3 was used. In both cases, the gas was introduced from room temperature to the growth temperature of 620°C.

One can remark that MOVPE regrowth introduce a smoothing of the grating shape but no change in the depth whatever the PH_3 partial pressure is. The deformation of the grating profile seems to be a consequence of mass-transport caused by the presence of PH_3 as explained in [3]. But, on the other hand, PH_3 can be useful to prevent the departure of phosphorus from the surface and so to minimize the non-radiative recombination of the injected carrier at the InP-InGaAsP corrugated interface. So we better use 1.2% PH_3 in the carrier gas while the deformation is equivalent in both case.

3. LASER STRUCTURE AND FABRICATION

The structure, sketched in figure 3, is fabricated as follows : first, a n-type InP buffer layer, an undoped 1.5 µm GaInAsP active layer (0.2 µm thick) and a p-doped 1.3 µm GaInAsP guide layer (0.2 µm thick) are

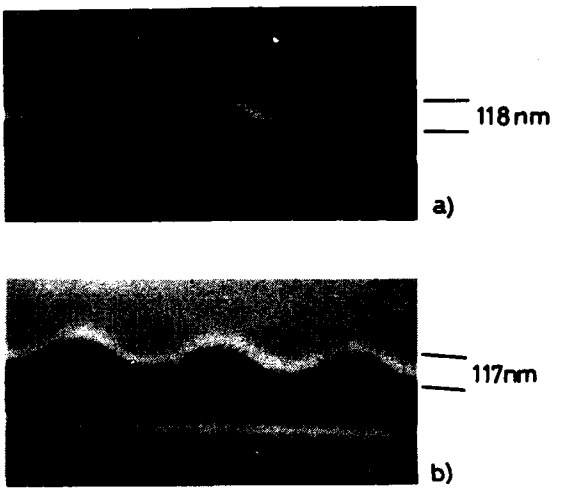


FIGURE 2
MOVPE growth over grating
a) with 1.2% PH_3 in the carrier gas
b) with 0.12% PH_3 in the carrier gas

successively grown by LPE.

After photoluminescence characterisation and epilayer thicknesses measurement, the grating parameters -period and shape- are determined. The grating is made by conventionnal holographic exposure, the photoresist width is controlled by HeNe (λ= 6328 Å) diffraction efficiency measurement. The corrugation is chemically etched into the waveguide layer using SBW solution [8], depth and shape estimations can be done again from HeNe diffraction efficiency measurement.

Then a stripe is defined into the quaternary layers by classical lithography and selective

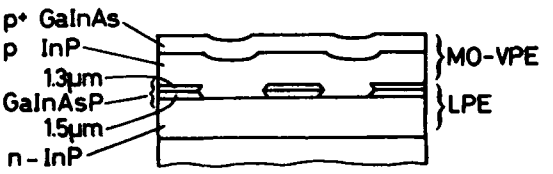


FIGURE 3
Schematic view of the BH-DFB structure grown by hybrid LPE-MOVPE.

chemical etching through an Si_3N_4 mask. Figure 4a shows the aspect of a 2 µm wide stripe, one can notice the grating lines perpendicular to the stripe direction.

In a second epitaxial step, a p-type InP cladding layer and a p^+-type InGaAsP contact layer are grown by atmospheric pressure MOVPE. The pregrowth conditions are those presented above. The preservation of surface morphology is also efficient for the stripe as can be seen on the SEM photography of the Buried-Ridge-Stripe (BRS) structure (figure 4b).

After the regrowth, the definition of a new 5 µm wide stripe in the contact layer, associated with an oxyde layer on each side, allows the localisation of the current injection through an ohmic p-contact. The wafer is then thinned down to about 80 µm and back-side metallisation is done for n-contact. Finally, the wafer is cleaved to obtain 250 µm-long laser diodes .

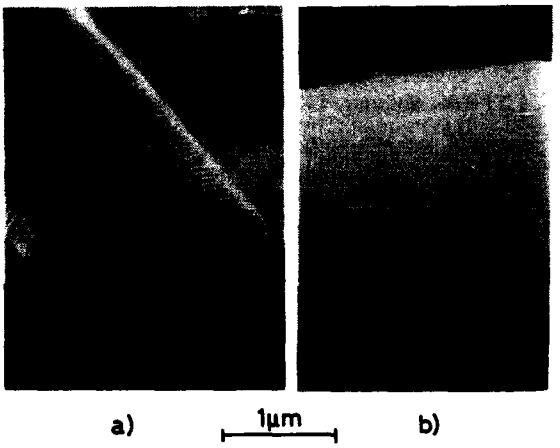


FIGURE 4
Etched stripe a)before and b)after MOVPE regrowth

4. LASER CHARACTERISATION

The current confinement of such a laser was described several years ago for structure grown by LPE [9] and later by Low-Pressure MOVPE [10]. The enhancement of the current flow through the buried ridge is due to the difference in diffusion potentials between the p-n InP homojunction on each side of the stripe and the p-n InP/InGaAsP heterojunction within the stripe. In our case today, the good quality of the p-n InP homojonction fabricated by hybrid epitaxy process leads to an effective current confinement up to 150 mA.

As cleaved, the pulsed L-I curve of such a laser is shown on figure 5. Threshold currents as low as 38 mA are obtained. This value is quite low in spite of the use of a λ=1.3µm guide layer which leads to a poor electron confinement in the active region [11]. Better confinement and lower threshold current could be obtained with lower band-gap guide layer.

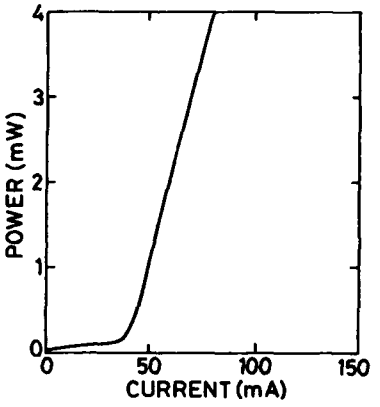


FIGURE 5
Pulsed L-I curve with 38 mA threshold current

For CW operations, lasers are mounted on copper heat sinks. In order to ensure a single mode operation of most of the diodes, one cleaved facet is anti-reflection coated. This prevents the competition between Fabry-Perot (FP) and DFB propagation modes as well as the two-mode behaviour theoretically predicted [12] for a perfectly symetrical cavity.
After this treatment, the discrimination between lasing mode and the most important side mode is over 30 dB.
Figure 6 shows a subthreshold optical spectrum where a 23.7 Å wide stop-band is clearly visible.

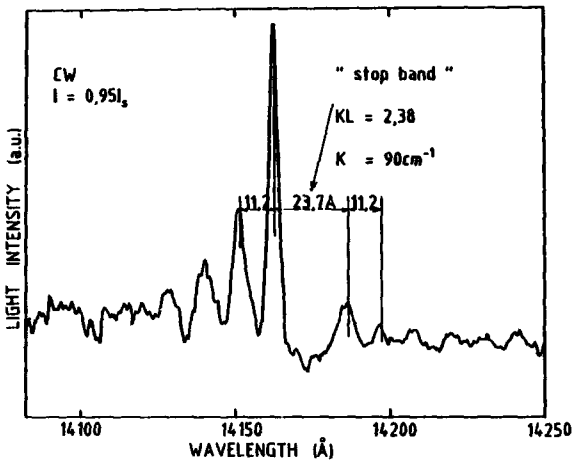


FIGURE 6

DFB spectrum below threshold

Using the coupled-mode theory[12], we deduced from this measure a coupling coefficient of 90 cm^{-1} (KL = 2.3).

Single-longitudinal mode operation has been maintained up to 8 mW as can be seen on the optical spectra in insert (figure 7). Higher power is limited in this structure by current leakage through the InP homojunction.

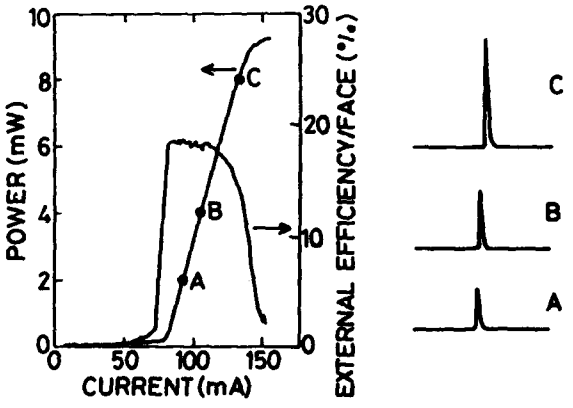


FIGURE 7

CW L-I curve and optical spectrum

5. CONCLUSION

We presented a process for the fabrication of BH-DFB lasers including MOVPE regrowth over the grating. Threshold current down to 38 mA and single mode operation up to 8 mW have been obtained. A coupling coefficient of 90 cm^{-1} has been deduced, demonstrating the ability of MOVPE to prevent any significant grating deformation during the regrowth.

ACKNOLEDGEMENTS

We should like to thank M. Carré and D. Robein for technical assistance.

D. Lesterlin thanks the CNRS for financial support through a Docteur-ingénieur scholarship.

REFERENCES

[1] M. Razeghi, B. De Crémoux, J.P. Duchemin
J. of Crystal Growth 68(1984)389-397

[2] A. Mircea, R. Mellet, B. Rose, P. Daste, G. Schiavini
J. of Crystal Growth(77)340-346, 1986

[3] A.W. Nelson, L.D. Westbrook, P.J. Fiddyment
IEE Proceedings, vol.132, Pt.J, n.1, Feb.85

[4] C.J. Armistead, B.R. Butler, S.J. Clements, A.J. Collar, D.J. Moule, S.A. Wheele
Electronics Letters 23(11), p.592, 1987

[5] Y. Itaya, M. Oishi, M. Nakao, K. Sato, Y. Kondo, Y. Imamura
Electronics Letters 23(5), p.193, 1987

[6] M. Oishi, M. Nakao, Y. Itaya, Y. Imamura
10th IEEE Int. SC. Lasers Conf. JAPAN 86

[7] J. Kinoshita, H. Okuda, Y. Uematsu
Electronics Letters 19(6), 1983

[8] T. Saitoh, O. Mikami, H. Nakagome
Electronics letters 18(10), 1982

[9] J.C. Bouley, J. Charil, G. Chaminant
9th IEEE int. SC Laser conf. BRAZIL 84

[10] R. Blondeau, M. Razeghi, M. Krakowski, G. Vilain, B. De Crémoux, J.P. Duchemin
Electronics Letters 20(21), p.850, 1984

[11] L.D. Westbrook, A.W. Nelson
J. Appl. Phys. 56(3), p.699, 1984

[12] W. Streifer, R.D. Burnham, D.R. Scifres
IEEE QE-11, n.4, p.154, 1975

IP.10. - INVITED PAPER

LIGHT-GUIDED ETCHING FOR III-V SEMICONDUCTOR DEVICE FABRICATION

Dragan V. PODLESNIK

Microelectronics Sciences Laboratories and Center for Telecommunicatons Research, Columbia University, New York City New York 10027, U.S.A.

The rapid, ultraviolet-induced aqueous etching produces vertical, high aspect features in GaAs samples of different crystal orientations. Much of the speed and anisotropy of the etching is attributed to the formation of efficient hollow, optical waveguides. These guides have been characterized by measuring the optical loss and the field distribution within the guide. The optical loss is typically small and does not restrict the etching of deep features.

1 INTRODUCTION

Rate anisotropy, which results in a strong spatial directionality, is an important characteristic in many semiconductor processing operations. In recent years, the work with the anisotropic etching has focused on its use for the machining of semiconductor materials. Vertical, high-aspect features are an important requirement in the fabrication of advanced electronic and micromechanical devices[1]. The close spacings and small sizes of modern integrated circuits require vertical etching to eliminate the undercutting of adjacent structures. In addition, the third dimension of a semiconductor wafer is currently being investigated as a means of providing additional electrical interconnections[2]. In particular, the interest in these through-wafer interconnections stems from three basic goals: to reduce the length of connections between devices for faster processing rates; to reduce the interference and crosstalk between interconnections; and to reduce the area occupied by interconnections on the surface of the semiconductor wafer. An example is a GaAs FET with via connections through the substrate[3]. Such transistor has higher gain at

microwave frequencies, higher power density, and potentially lower manufacturing costs than a conventional GaAs field effect transistor.

Finally, anisotropic etching has been extensively used in nonelectronic applications[1]. Even simple holes and grooves etched in a semiconductor wafer (typically silicon) can be utilized for many applications. One usage is the generation of high precision molds for microminiature structures. Patterns etched clear through the wafer can be applied in the area of ink jet printing technology. In particular, the hole on the bottom of the wafer is used as an orifice, typically about 20 um, for an ink jet stream[1].

2 MECHANISMS FOR ANISOTROPY IN ETCHING

Conventionally, etching anisotropy in single-crystal materials is achieved by relying on a crystal plane dependent process, such as reaction rates in wet chemical etching. This crystallographically sensitive etching has been successfully used to produce deep vertical features, but its utility is restricted by the requirement of specific and limited crystal orientations. Localized electrochemical jet etching has been also used to generate vertical holes in semiconductor wafers.

When etching occurs by an ion assisted reaction, etch rate anisotropy can be expected, because ions are incident normal to the wafer surface. When a semiconductor is plasma etched, incorporation of an appropriate gas additive results in the formation of a passivating film that prevents side wall etching. However, a perfectly vertical, through-wafer feature cannot be made with ion and electron beam sources because of the spreading of the particle beam and loss of the particle fluence with increasing feature depth. In addition, the massive particle bombardment produces typically an incurring damage of the semiconductor surface.

In this paper we will show that laser-induced aqueous etching with the ultraviolet beam can be used to make deep, high-aspect features irrespective of crystal orientation. If the laser beam is incident normal to the semiconductor surface, it could be expected that the light-assisted etching will occur primarily in the direction of the beam. But, unlike ion or electron beams, the laser beam is guided by the etched structure itself via glancing-angle reflections thus resulting in the vertical etching. This waveguiding effect coupled with the rapid etching at low laser intensities was first seen in the ultraviolet (257 nm) induced aqueous etching of GaAs[4].

This anisotropic, laser-directed etching is then also seen in other semiconducting materials[6] and in some gas phase, laser assisted process[7]. In all cases, the etching is independent of crystal orientation, and does not rely on bombardment with massive particles. Furthermore, rapid etching is already achieved at low laser intensities, thus laser-induced damage of the semiconductor surface, which is characteristic for high intensity irradiation, can be avoided.

3 EXPERIMENT

For the etching experiments in the ultraviolet, the 257 nm output from a frequency doubled argon-ion laser was used. The laser light was manipulated by a scannable optical microscope; using a computer-controlled microscope stage, we could scan the sample perpendicularly to the axis of the laser beam. The laser beam was focused to a 3-4 um spot with a 10x microscope objective (N.A. = 0.2) onto the sample surface. The semiconductor sample was mounted inside a quartz cell, which was filled with a liquid. For this optical configuration a measured confocal beam parameter was approximately 25 um. The incident laser power was measured with an ultraviolet photodiode placed underneath the microscope objective. The power ranged from 10 nW to 100 uW, corresponding to laser power densities from 100 mW/cm^2 to 1 kW/cm^2; typically, 1 W/cm^2 was used in experiments.

Most of our experiments were performed on single-crystalline (100) GaAs, doped with Si ($n=10^{18}$ cm^{-3}). The samples were of different thicknesses which ranged from 40 um to 250 um. In addition, three different crystallographic orientations of GaAs crystals, (111)A, (111)B, and (110), were briefly examined for the sake of comparison. The etchant was a diluted HNO_3 aqueous solution (1:20 by volume). This solution causes no dark etching and only weakly absorbs ultraviolet light. The solution, in conjunction with the ultraviolet illumination, allows efficient etching of all doping types of GaAs.

4 FORMATION OF HIGH ASPECT SURFACE RELIEFS

The ultraviolet optical properties, i.e. the high, near metallic reflectivity[7] give rise to an unexpected effect seen in material removal reactions on GaAs surfaces. This effect is the extremely high-aspect features that arise as a result of optical waveguiding in surface-relief structures as they develop on the surface. This waveguiding effect, coupled with

the rapid etching at low laser intensities, was first seen in the ultraviolet induced aqueous etching of GaAs. Subsequent experiments have shown it to be a more general phenomenon, that is applicable to other semiconducting materials, e.g., Si and InP. Light-guided feature formation is observed when a focused, patterned, or scanned ultraviolet laser beam is used. In all cases in the ultraviolet, the etched structure confines and efficiently transmits the laser beam.

The possibility that the vertical, ultraviolet-induced etching relies on the semiconductor crystal anisotropy was ruled out after the etching experiments with GaAs samples of different crystallographic orientations. A weak crystallographic dependence is found, resulting in average etch rates of 7:9:9:11 um/min for the (111)A, (100), (110), and (111)B faces, respectively. However, perfectly vertical features are obtained for all the etchings, proving that the crystal orientation does not affect the etched profiles. This characteristic is important since it allows maskless etching of vertical, high-aspect profiles in all crystallographic directions. Note that, whereas the ultraviolet anisotropic removal is virtually independent of crystallographic orientation of the semiconductor crystal, visible light does show a somewhat stronger crystallographic dependence, as reported in [8].

Figures 1 to 2 show typical profiles of microstructures such as via holes and slits formed through a GaAs wafer. The slits were etched by scanning the wafer perpendicular to the axis of the laser beam. In all experi-

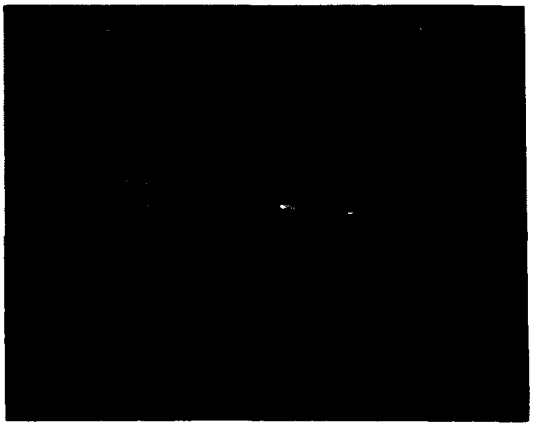


Fig.1 SEM micrograph of slits etched through a GaAs wafer.

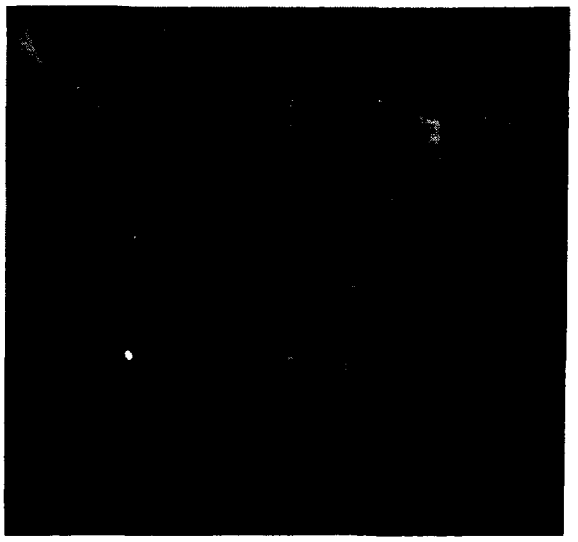


Fig.2 SEM micrograph of a etched groove in GaAs: close-up of the side wall.

ments, the laser beam was focused such that the beam waist was on the front of the semiconductor wafer. The entrance of the structures is well defined and the surrounding area is undisturbed. A distinct characteristic of these features is their smooth and vertical side walls. Their width, typically 2-3 um, remains virtually constant, independent of the etched depth. In contrast to this, the focused laser beam, with the 3 um waist diverges considerably over the corresponding distance, as shown in Fig. 3 . The measured confocal beam parameter is only 25 um as compared, for example, to the perfectly vertical etch through the 200-um-thick wafer. This clearly shows that the processing beam is confined inside the hollow, etched structure which prevents the beam from diverging.

Because the vertical features are created by using a Gaussian beam, one might think that their profiles should also appear Gaussian. The vertical walls is a strong deviation from the Gaussian shape. In order to determine a mechanism for the formation of vertical waveguides, we monitored the development of holes, as shown in Fig.4. Initially, the etch profile is essentially identical to that of the incident laser beam, as shown in Fig. 5. However, as the etch depth increases, the feature assumes a tubular, non-Gaussian profile, Fig.5

The initial formation of these anisotropic features from a focused

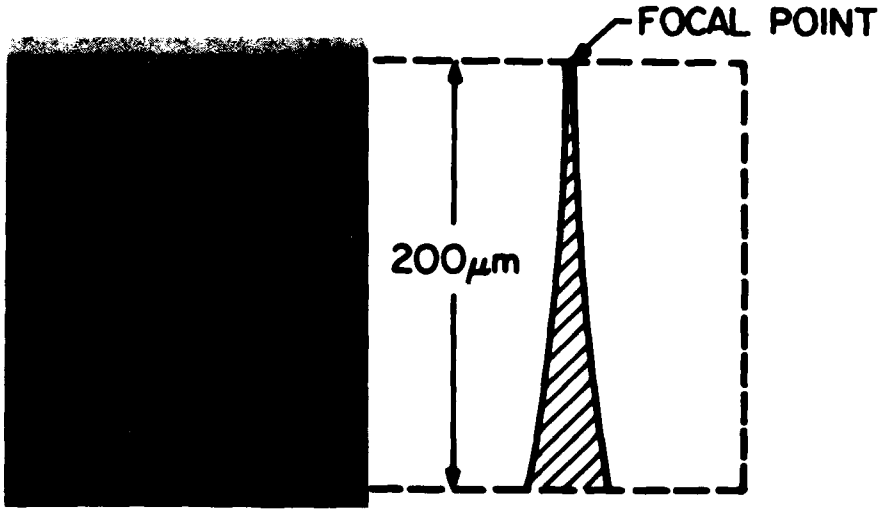


Fig.3 Comparison of the etched profile with free-space propagation of the focused laser beam.

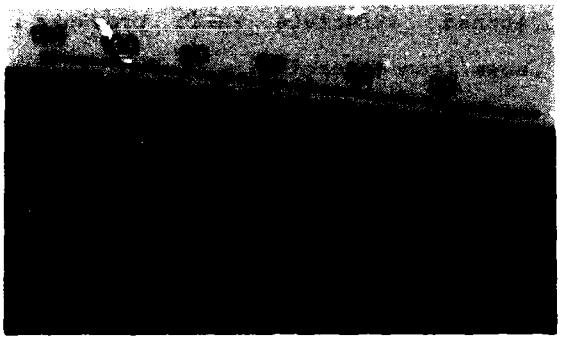


Fig.4 SEM micrograph of the temporal development of the vertical via hole.



Fig.5 Initial, Gaussian-shape etchings and the formation of tubular, non Gaussian profiles.

laser beam with a beam waist of 4 um can be calculated with a model based on the use of ray optics. The model includes the very important angular dependence of reflectivity and absorption, which for the case of a lossy, metal like material must be obtained from a modified version of the Fresnel equations. In this case, the surface slope, which is equal to the local incident angle of the laser beam, is given by

$$\theta(z,r,t)=\tan^{-1}[dz(t)/dr(t)] \quad (1)$$

where z(t) and r(t) are, respectively, the time-dependent vertical and radial coordinates of the surface structure. The two components of the local interface velocity of removal are given by

$$(dz/dt)=\{kI(r)(1-R(\theta))\cos(\theta)\}\cos(\theta) \quad (2)$$

$$(dr/dt)=\{kI(r)(1-R(\theta))\cos(\theta)\}\sin(\theta) \quad (3)$$

where I(r) represents the incident Gaussian beam and (1 - R(θ)) is the local absorption. The interface velocity, in braces, is assumed to be linearly proportional, by the reaction rate constant k, to the local absorbed light intensity. The cos(θ) term in Eqs. (2) and (3) represents the change in incident power deposited on the surface due to surface tilt. The last term in Eqs. (2) and (3) represents the directionality of etching with this surface tilt. Material removal is always perpendicular to the plane of the surface.

During the very initial stages of formation of, say, a hole, with a Gaussian laser beam, all the angular dependent terms are inconsequential. As a result, initially the surface should follow the intensity distribution of the laser beam. Therefore the initial feature will be Gaussian.

As the feature walls become more vertical, the angular-dependent terms mentioned earlier become more important. Even then, for an incident Gaussian laser beam, the surface structure

takes a non-Gaussian shape. However, despite this non-Gaussian effect, the above model will not produce the striking features with vertical walls seen here. Many of these effects occur because light begins to be trapped inside and undergoes internal reflections as the hole deepens. Owing to the angle dependent reflectivity, there is still enough power reflected off the sidewalls to be subsequently absorbed as the light again strikes a different portion of the surface. This first internal reflection, or second strike, does contribute significantly to the formation of the features. Figure 6 shows this computer simulation of via hole development that includes internal reflections to the first order. These

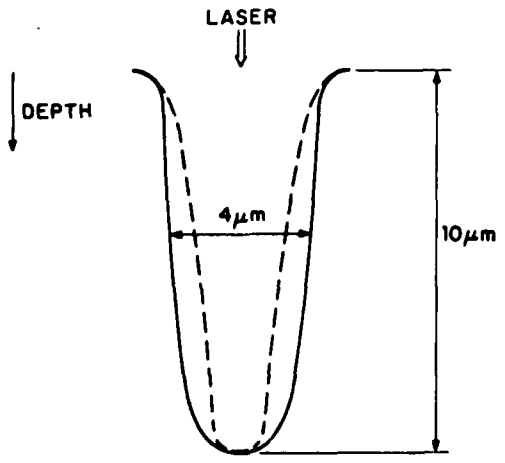


Fig.6 Calculated hole development with (solid line) and without (dashed line) first internal reflection.

reflections can have a pronounced impact on the developing shape of a cylindrical hole.

All the model figures display an overall appearance that is in agreement with an actual via. Eventually, a portion of the side wall becomes sloped to such a glancing angle that the laser light from the first strike is completely reflected. Also,when the hole is deep enough, the second strike resulting from the reflections off the bottom does not affect the shallower parts of the walls. Wall erosion therefore eventually stops. The incident beam is then optically guided through the hole, while vertical etching continues at the bottom. Under these conditions, etching can bee properly modeled as a self propagating waveguide.

In fact, much of the speed and anisotropy of the etching can be attributed to the properties of these optical waveguide-like features. In particular, the optical loss is small, even for 200- to 300-um-thick samples, and does not significantly limit the etch rate with increasing feature depth. Further, the continuation of etching even in the heavily solvated liquid at the bottom of a deeply etched feature is indicative of the active chemistry induced by the ultraviolet light.

5 OPTICAL WAVEGUIDE FORMATION

One of the most interesting consequences of the unusual laser chemistry described above is that it permits the formation of long cylindrical tubes or vias in semiconductor wafers. In fact, they are miniature, hollow semiconductor waveguides. This unusual microstructure has suggested a number of novel applications in microelectronics and integrated optics, including vertical, high-density optical fiber interconnects,[9] through-wafer optical light guides,[10] and high density, through-wafer vias for electrical interconnections from the front to the back of silicon or GaAs integrated circuits[3].

It was already suggested that the guiding of the laser beam is an important factor in determining via-hole profiles and etch rates. The circular hollow dielectric and metallic waveguides have been thoroughly discussed by Marcatili and Schmeltzer[11]. To illustrate some of the features of the hollow semiconductor waveguide, we applied a similar analysis for the via holes etched through GaAs samples. The theory and the experiment showed excellent agreement. Figure 7 shows the attenuation coefficient of the hollow GaAs waveguide as a function of hole diameter. In agreement with the theory, the measured attenuation coefficient was 6.3 times larger for 514 nm than for 257 nm light. Such high optical attenuation in conjunction with less efficient photochemistry is a reason that the visible laser light does not exhibit the same waveguiding effect.

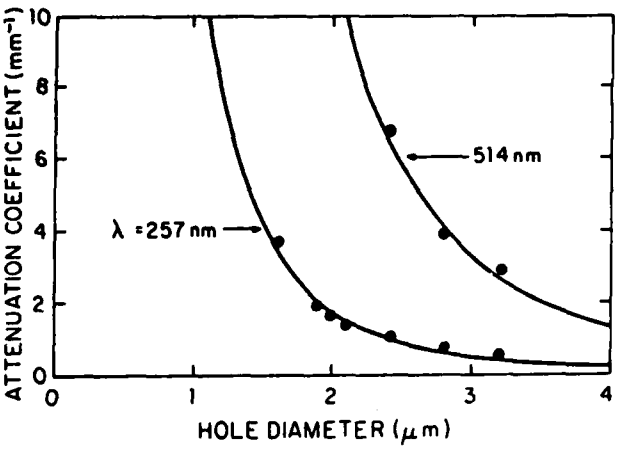


Fig.7 Attentuation coefficient of the hollow GaAs waveguide as a function of hole diameter, measured for 257-nm and 514-nm wavelengths.

6 CONCLUSIONS

In this paper we have shown that the rapid, ultraviolet-induced process produces high-quality vertical etching through standard thickness GaAs wafers. This directional etching is attributed in part to the formation of efficient optical guides, so that the laser processing-beam is guided by the etched structure itself. Applications of this

technique to novel electro-optical devices and through-wafer electrical connections are currently being explored.

At this point, however, many important questions remain unanswered. For example, in the very deep etched features, to what degree does capillary flow caused by concentration or density gradients influence the etching? Furthermore, in the formation of submicrometer gratings by laser interferometry, it is known that surface electric fields formed between dark and illuminated areas significantly influence the shape of the grating groove profiles. To what degree do surface fields influence the etched features seen in our experiments? These questions as well as the nature of the microscopic chemistry at the semiconductor surface remain fertile ground for further studies.

ACKNOWLEDGEMENT

The author would like to thank A. Willner, H. Gilgen, and R. Osgood for many useful comments and suggestions. Portions of this research were supported by the defense Advanced Research Projects Agency, the U.S. Air Force Office of Scientific Research, the U.S. Army Research Office, and the U.S. Joint Services Electronics Program.

REFERENCES

[1] K.E. Petersen, Proc. IEEE, 70, 420 (1982).

[2] T. Morie, K. Minegheshi, and S. Nakajima, IEEE Electron Desvice Lett., EDL-24, 411 (1983).

[3] L.A. D'Asaro, J.V. DiLorenzo, and H. Fuki, IEEE Trans. Electron Devices ED-25, 1218 (1978).

[4] D.V. Podlesnik, H.H. Gilgen, and R.M. Osgood, Appl. Phys. Lett., 45, 563 (1984).

[5] D.V. Podlesnik, T. Cacouris, H.H. Gilgen, and R.M. Osgood, Jr., to be published in Appl. Phys. Lett.

[6] C. Arnone, M. Rothschild, and D.J. Ehrlich, Appl. Phys. Lett., 48, 736 (1986).

[7] H.R. Philipp and H. Ehrenreich, Phys.Rev. 129, 1550 (1963).

[8] R.M. Osgood, A. Sanchez-Rubio, D.J. Ehrlich, and V. Daneu, Appl. Phys. Lett. 40, 391 (1982).

[9] P.R. Prucnal, E.R. Fossum, and R.M. Osgood, Opt. Lett., in press.

[10] D.V. Podlesnik, H.H. Gilgen, and R.M. Osgood, Appl. Phys. Lett. 48, 323 (1986).

[11] E.A.J. Marcatili and M.A. Schmeltzer, Bell System Tech. J., 44, 1783 (1964).

IP.11. - INVITED PAPER

DESIGN AND PERFORMANCE OF HIGH POWER SEMICONDUCTOR LASERS

K. Mettler

Siemens AG, Forschungslaboratorien, Otto-Hahn-Ring 6, D-8000 München 83, FRG

A discussion of fundamental power limits of GaAlAs/GaAs and InGaAsP/InP lasers is followed by a review of design approaches and of the current status of development of high-power semiconductor lasers.
Newly designed structures and sophisticated crystal growth techniques have shifted the power limits of coherently emitting GaAlAs single-element lasers and arrays beyond 1 W of continuous wave (CW)optical power, while InGaAsP lasers of comparable structures have so far been thermally limited to about 200 mW CW power output. Power levels assuring adequate device reliability of around 100 mW are predicted for lasers from both material systems. Thus the application potential of semiconductor diode lasers has been appreciably widened.

1. INTRODUCTION

Due to their many attractive features, semiconductor diode lasers are especially fascinating among the solid-state devices of today. They offer distinct advantages over gas and other lasers: small size direct modulation capability, high power conversion efficiency, low drive voltage and high reliability.

The realization of these advantages in a practical device under the increasing demands of various applications has remained a challenging task over the past 25 years since the first demonstration of the laser diode. The complex and intricate fabrication technology required by these devices was responsible for the enormous research and development effort expended by many laboratories throughout the industrialized world which led to the present stage of development where a multitude of laser structures are availble for many different tasks. The important industrial applications of laser diodes arrived first in the late 1970s in fibre-optic communications and in digital audio-disk players; intense research and development efforts together with very rigorous system demands finally resulted in a first generation of compact, rugged and reliable devices which had a stable beam, modulation capability of up to 1Gbit/s, and could be efficiently coupled to optical fibres. A rather low optical power output of several milliwatts was mostly sufficient for these applications.

In recent years the development of semiconductor lasers is characterized by increasing efforts to extend performance limits which is demanded by many novel applications. Since it became evident that the achievable optical output power could be dramatically increased beyond the low milliwatt range, there is vital interest to exploit this high-power capability for optical recording systems, laser printers, data distribution systems, space-communication between satellites, pumping of miniature solid-state lasers, micro-surgery and other applications. An eventual success of the diode lasers - in competition with the gas laser - depends much on the extent to which it can meet the additional technical demands imposed by these applications, particulary as far as optical power and related features are concerned. This paper summarizes basic limitations, design principles and solutions obtained over the past few years towards achieving maximum optical power output from semiconductor lasers.

2. REQUIREMENTS AND DESIGN ASPECTS

The above-mentioned applications impose new stringent requirements in addition to those known for low-power laser diodes. Most important are a high output power level (typically beyond bout 50 mW) as well as a narrow far-field pattern and - in specific cases - a small spectral width. With the fulfillment of these demands, the diode laser has become one of the most highly stressed semiconductor devices in terms of intense optical and electrical fields and thermal loading. Thus one must pay special attention to the question of a satisfactory lifetime. Consequently, a basic goal in the development of diode lasers is to extend their catastrophic and thermal limits to very high optical power (Fig. 1) so that somewhat lower usable power levels are then achieved with negligible degradation of perfor mance. The lasers must therefore be designed to have a high overall efficiency for the conversion of electrical to optical power which means that particularly high slope efficiencies η_d and and low threshold currents I_{th} are required in addition to a good thermal design of the laser chip and heat-sink (Fig. 2). These aspects and other demands for diode lasers lead to a multitude

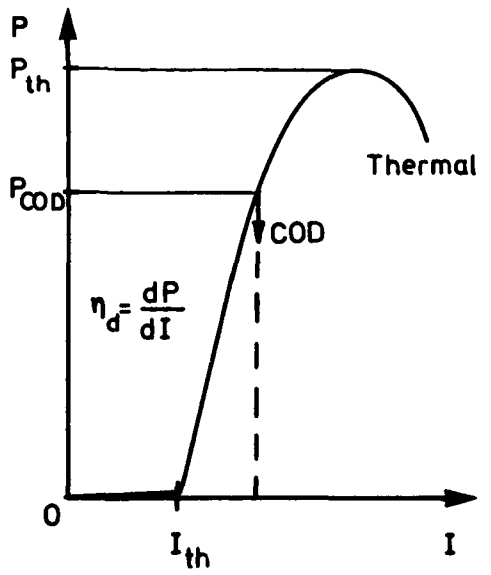


FIGURE 1
Light-current characteristics of a semiconductor diode laser, in a schematic representation: I diode current, I_{th} threshold current, P optical power, P_{COD} power limit due to catastrophic optical degradation, P_{th} power limit due to thermal effects, η_d differential efficiency.

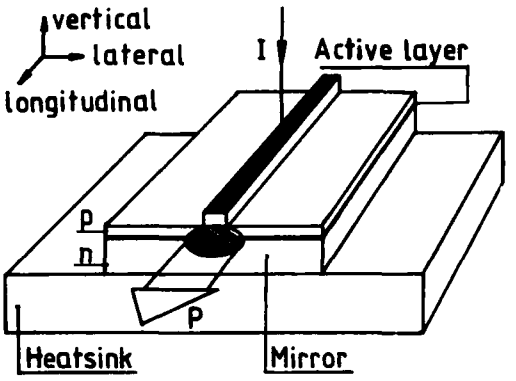


FIGURE 2
Schematic representation of a semiconductor diode laser.

of complex and sometimes even conflicting requirements. Solutions have been found through newly developed laser structures which are different from and more sophisticated than the earlier low-power lasers. These either allow for additional degrees of freedom in design or inherently rely on an improved control of structural parameters through the utilization of precise material growth methods such as MOVPE or MBE. In the following, a more detailed systematic description of solutions for obtaining high optical power from semiconductor lasers is given together with the features of some recent diode structures representative of the state of the art in this field.

Short-wavelength devices - composed of GaAlAs/GaAs heterostructure layers and emitting at wavelengths between 0.77 and 0.88 µm - and long-wavelength devices - made of layers

of InGaAsP/InP and radiating at 1.3 to 1.6 µm - will be treated in separate sections since maximum output power is limited by different mechanisms im the two material systems.

3. GaAlAs/GaAs LASERS

Output powers of conventional GaAlAs lasers under pulsed and CW operation are limited by catastrophic optical damage (COD) at the laser mirrors /1/. Additionally, erosion of unprotected facets by the atmosphere may occur even at much lower optical power levels /2/. The basic mechanism responsible for COD is the local heating caused by absorption of radiation near the facets /3/ and by nonradiative recombination of carriers at the facets /1/.

For pulsed laser operation, the catastrophic optical power density limit (P_{COD}) depends on the time duration of the optical flux and on the surface recombination velocity s at the facets (Fig. 3).

Under CW operation P_{COD} is given in good aproximation /4/ by

$$P_{COD} < \frac{1-R}{1+R} \cdot \frac{1}{\Gamma} \cdot \frac{1+\frac{D}{sL}}{\alpha L} \cdot q_{max}$$

where R is the mirror reflectivity, Γ the optical confinement factor, α and L are respectively the effective absorption coefficient and diffusion length of the carriers in the absorbing region at the laser facet, and q_{max} the maximal heat flux at which the facet ist destroyed.

For ordinary GaAlAs/GaAs DH laser structures D, L, s and q_{max} are specific material parameters which cannot be significantly improved; q_{max} has been estimated to be around 3 MW/cm^2 /4/, giving experimentally confirmed values of the critical external optical power density of up to 1 MW/cm^2 /15/ depending on active layer thickness and Al contrations. To obtain larger absolute optical

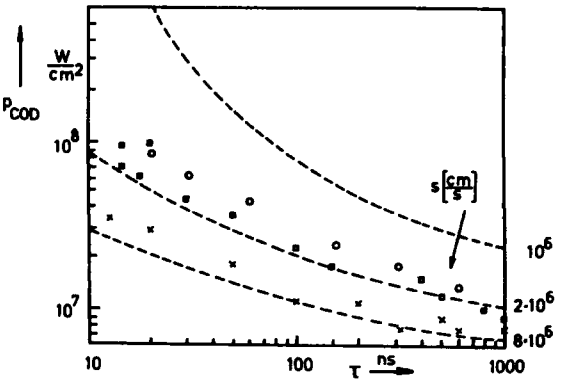


FIGURE 3
Optical power density limit P_{COD} of catastrophic optical mirror damage as a function of pulse width τ for GaAlAs/GaAs oxide stripe lasers with different methods of preparation (from /4/): LPE-grown lasers (x) with 1 kV sputtered mirrors, (□) with 300 V sputtered and Al_2O_3-coated mirrors, (○) MO VPE-grown lasers with cleaved mirrors. Dashed curves are calculated for 0.1 µm active layer thickness and 3 µm stripe width; the surface recombination velocity s at the facet is varied as a parameter.

power values, various ways have been used singly or in combination, to influence the remaining parameters α, R and the spot size, which has resulted in a diversity of laser structures. Early attempts were aimed at decreasing the ratio of inner to outer optical power through the use of anti-reflection coatings on the facets. Today, this is still used in addition to one of the following more basic methods:

Decreasing local absorption and carrier flux: A very effective approach to eliminate the effect of local absorption of laser radiation and the resulting carrier flux to the facet is the non-absorbing mirror (NAM) or "window" laser. In this solution, regions are created adjacent to the mirrors which are transparent to the laser radiation coming from the inner part of the resonator. A NAM effect is achieved either by selectively increasing the band gap of the material at the facet or by lowering the band gap and the

photon energy of the laser radiation in the inner region of the active layer. This principle was first demonstrated for the latter case: by diffusion of zinc into the laser cavity except at the facets, the lasing energy was shifted to a long wavelength. A markedly increased CW optical power density of up to 4.7 MW/cm^2 and stable fundamental mode operation up to 55 mW CW were thus obtained /5/. Such a laser type still has a drawback, however: it requires a precisely controlled diffusion. Lasers with a transparent higher band gap facet have been realized in different ways:

(i) In buried-facet type structures a part of the facet material was etched away during the fabrication process and replaced by a higher band gap material by regrowth /6,7/. Forcing the light to propagate in a non-absorbing LOC waveguide (cf.following section) up to the facet plane - in addition to the previously mentioned diffusion - a similar power density as for the diffused laser and CW output powers of up to 88 mW have been attained /7/.

(ii) Impurity-induced disordering (IID) has been shown to be a highly promising technique to increase the band gap of quantum well materials e.g. selectively disordering by diffusion the regions near the mirrors /8/.Of the possible impurities Zn and Si, the latter is preferable /9/ due to the significantly lower concentrations needed as compared to Zn where the higher required concentrations result in high free carrier absorption. An increase in the ultimate optical power of multiquantum well (MQW) lasers by a factor of 2.5 is achieved as compared to MQW lasers without any facet protection /10/.

Introducing current-blocking layers below the active layer only at the cavity ends has been used as another approach to suppress local temperature rise near the facets. In combination with an enlargement of the near field spot size (cf. following section) and application of reflective and anti-reflective coatings respectvely to the rear and front facets, a COD power density of about 5 MW/cm^2 in CW and an absolute optical power of 120 mW could be obtained at an emission wavelength of 780 nm /11/.

Increasing spot size:

A successful approach to high power diode lasers consists of enlarging the spot size of the near-field pattern to reduce the optical flux density by structuring the active area both vertical by and parallel to the junction plane (Fig. 2) as illustrated schematically in Fig. 4. For the vertical increase of the spot a frequently employed approach is the TAL structure /12/: A thin active layer (TAL) with a thickness below 0.1 µm acts as a very weak waveguide, so that as much as 90 percent of the optical energy spreads out into the (non-absorbing) cladding layers. Alternatively, the large optical cavity (LOC) structure /13/ broadens the effective area of the near-field spot with the aid of a relatively thick waveguide layer adjacent to the active layer. In comparison with the TAL structure the LOC structure has a disadvantage because it requires the precise control of two parameters, i.e. the thickness and the composition of the guide layer to avoid poor mode stability in the direction perpendicular to the junction. In the TAL structure the active layer thickness is the only critical parameter; this can now be well controlled through the well-established growth methods such as metal-organ ic vapour phasepitaxy (MOVPE) and molecular beam epitaxy (MBE). Even liquid-phase epitaxy (LPE) can bring forth an extremely thin active layer when specific mechanisms of growth on a ridged substrate are utilized. By the use of this concept together with an effective lateral confinement of the current and multilayer coating techniques for the facets, the buried twin-ridge substrate (BTRS) laser /14/ attains a maximum CW output power of 200 mW (Fig. 5) and a fundamental transverse mode

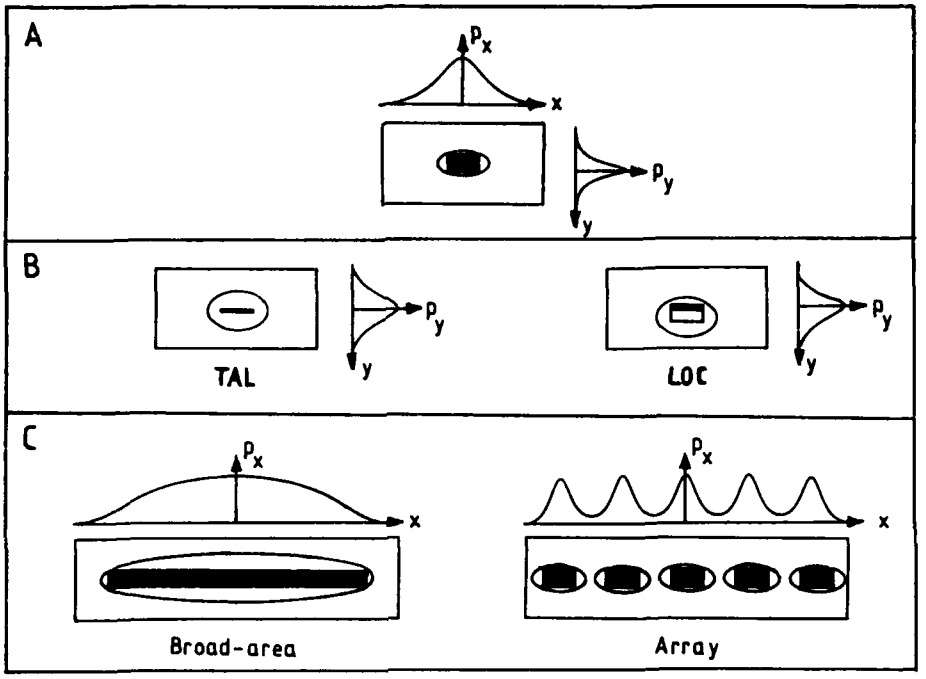


FIGURE 4
Schematic cross sections of diode laser structures developed to increase the size of the lasing spot at the mirror in the vertical (B) and lateral (C) direction with respect to the basic structure shown in (A). The ellipses around the active regions (dark) schematically represent the local field distribution; also given are the field intensities in the vertical (X) and lateral (Y) direction.

operation at more than 100 mW /14/. A decrease of the active layer thickness below about 0.05 µm leads, however, to a marked increase of threshold current density due to the decreasing mode confinement factor /15/. The threshold current density can still be lowered nevertheless through the use of a separate-confinement heterostructure (SCH) which confines the injected carriers to an even narrower layer and in which the light is separately defined in a wider optical cavity by outer heterojunctions /16,17/. In the special case of a graded-index separate-confinement heterostructure (GRIN-SCH), parabolically graded waveguide layers on both sides of the thin inner layer are used to support Gaussian-shaped waveguide modes /18/. SCH lasers have been recently produced employing a double quantum well (DQW SCH) as the inner vertical layer structure and a metal-clad ridge waveguide for the lateral confinement. These lasers combine high zero order mode output power (56 mW CW) with a low threshold current (10-12 mA) and a very high external differential quantum efficiency (41-46 %). The CW burn-off power density of these MOVPE-grown lasers of more than 6.5 MW/cm^2 (for uncoated facets) is the highest value reported to date /19/.

In addition to increasing the spot size in the vertical direction, it is also possible to spread the optical power in the plane of the junction (Fig. 4). A straightforward approach is to use a wide injection stripe. Although a broader cavity may permit the existence of higher order modes, this is not

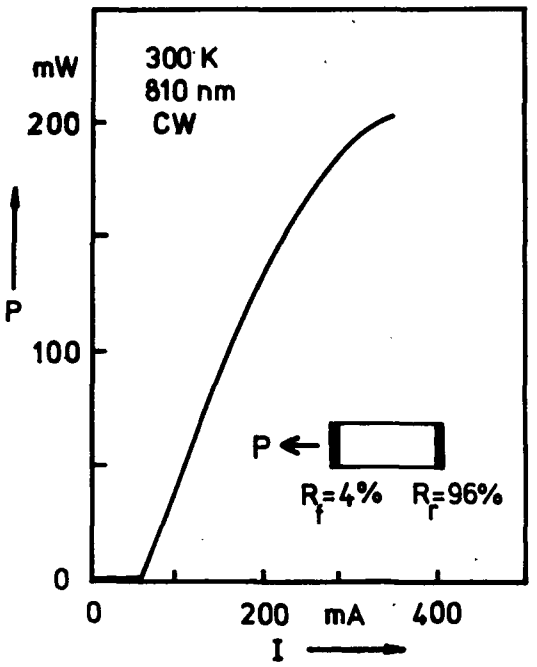


FIGURE 5
Output CW power P versus current I of a BTRS GaAlAs laser with TAL structure. R_f and R_r are the front and rear facet reflectivities, respectively (from /14/).

a principal disadvantage for those applications which require high optical power from the lasers but not necessarily concentrated in a single lateral mode. A serious drawback in early broad area lasers, were the slight inhomogeneities created through imperfect LPE layer growth and processing technology. The filamentation created thereby induced instabilities due to multilateral mode oscillation and caused the development of hot spots in the near-field pattern stressing the facet in that area. The very stringent requirements of high-performance broad area lasers with respect to lateral and longitudinal homogeneity could be satisfactorily met only recently with the advent of growth techniques like MOVPE and MBE. Thus e.g. broad-area GRIN SCH lasers with very thin active layers could be made by MBE giving excellent lasing uniformity across the entire width of a 200 µm wide facet /20/. A better homogeneity also improves the beam quality /21/: lasers with a 100 µm wide cavity and a single quantum well (SQW) GRIN-SCH structure grown by MBE exhibit a high degree of coherence along the junction plane and a stable and very narrow beam; in some cases as low as 0.8° FWHM. This is close to the diffraction limit and is of relevance for high-quality imaging. Another benefit of optimized growth conditions is an improved external quantum efficiency for which values of up to 80 % have been reported /22,23/.

The various structural approaches towards broad-area lasers fall into two categories. In the first category the applications require narrower waveguide structures (up to about 5µm wide) to maintain a single lateral mode /22,24,25/; in the second this condition has been eased. Successful examples for the latter case are a diode laser with a wider stripe giving 710 mW maximum CW optical output power /26/ (Fig. 6), 1 W output power been announced for a wider stripe /27/, and a 60 µm wide single-stripe GaAlAs SQW GRIN SCH laser has convered less than 3 W of electrical power to 1.5 W of multimode CW optical output /28/.

Demands to further increase the absolute CW output powers (i.e. much beyond 1 W) have led to the integration of semiconductor diode lasers into arrays (cf. Fig. 4).
Of the two types, i.e. incoherent and phase-locked arrays, the latter have the interesting property of achieving high optical power while maintaining a coherent optical laser output. The necessary lateral coherence can be achieved for up to a few tens of adjacent stripes limited by the present capability of epitaxial growth methods, array processing and chip-mounting techniques. These have to guarantee a sufficiently uniform lateral and longitudinal current injection and heat sinking.

First efforts in array development were aimed at a low threshold current density and

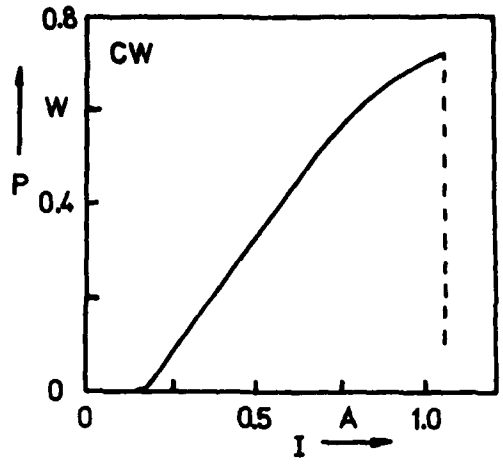


FIGURE 6
Optical CW power P versus current I of an MOVPE-grown broad-area GaAsAs laser (from /26/).

a high conversion efficiency as primary goals besides high power output. Using multi-quantum well and gain-guiding stripe structures for the vertical and lateral definition of the waveguide, CW optical powers in the vicinity of 2 W /31,32,33/ have been obtained (Fig.7). Beam quality was inferior due to the presence of higher-order spatial modes. Various schemes with sophisticated lateral and longitudinal structuring have been tried to promote fundamental mode operation and a narrow far-field pattern for the array. Diffraction-limited lobe with a FWHM of 0.7° has already been attained /34/.

Still higher optical powers have been obtained using a much higher number of stripes than is possible with coherent arrays. An increase in absolute power within one year could be attained starting with 5.4 W CW using individual arrays containing 100 to 140 stripes /35/ to 25 W for twenty arrays consisting of 40 stripes each arranged along a bar 1 cm long. They were driven in a quasi-CW mode with 150 µsec pulses /36/. Impressive values of 80 W /37/ and 114 W /38/ per centimeter of bar length have been recently announced.

Of importance for practical applications of semiconductor lasers are the maximum reliable optical powers. For GaAlAs lasers these are much below the record values. Yet single-element devices are about to extend the current range of reliable CW powers from 20 mW to 100 mW and above /39/. Thus, despite their high power density they are coming close to the reported power levels of arrays of 100 mW to 200 mW with median lives of 22000 and 4900 hours, respectively /29/.

4. InGaAsP/InP LASERS

Long-wavelength InGaAsP lasers are distinctly less prone to facet degradation than their GaAlAs/GaAs counterparts. On the other hand, InGaAsP lasers suffer from a much higher temperature sensitivity of to their threshold current density, so that their performance is seriously affected by thermal limitations. Therefore, one has to confine carriers and light to the lasing area particularly well in order to obtain high power output with low loss. This is achieved in present developments either by buried-heterostructure lasers incorporating leakage-current-blocking junctions beside the lasing stripe or by ridge-guide structures. Anti-reflection and high-reflection coating of the front and rear mirror, respectively, are also important. Optical powers of single-element 1.3 µm lasers of up to 150 mW and 200 mW CW have recently been reported respectively for distributed-feedback /40/ and Fabry-Perot type /41/ lasers. These record values are much below those of GaAlAs lasers. Higher powers, on the order of several hundreds of mW are obtained from InGaAsP edge- /42/ or surface-emitting arrays. Power levels approaching 1 W at 1.3 µm have been announced for the latter case.

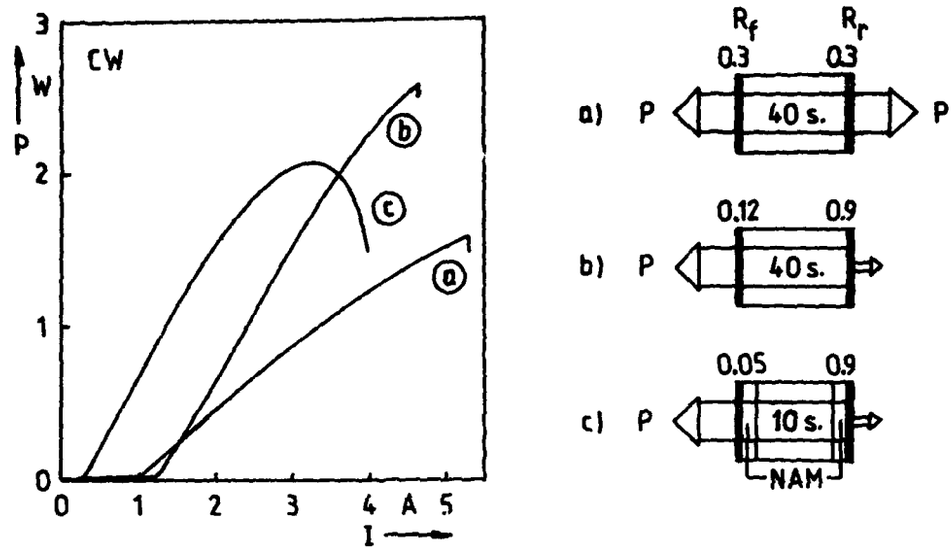


FIGURE 7
Light (P) vs. current (I) characteristics for three examples of GaAlAs/GaAs laser arrays with different structures:(a) 40 stripes, symmetrical mirror coatings /31/, (b) 40 stripes, asymmetrical mirror coatings /32/, (c) 10 stripes, NAM structure at both facets, asymmetrical mirror coatings /33/. R_f and R_r are the front and rear facet reflectivities, respectively.

5. CONCLUSION

Semiconductor lasers have already established a key role in a number of applications. With the newly detected possibility of achieving much more than 30 mW of reliable optical power and up to many watts of maximum power it will become increasingly attractive to substitute bulky, power-consuming and expensive light sources such as HeNe and other lasers or flash lamps with diode lasers in more and more applications. High-power semiconductor lasers have already demonstrated their usefulness in many areas even in the present status of their developement (cf. also /44,45/)and continuously challenge the R&D communityto extend their potential to an ever widening field of new applications.

ACKNOWLEDGEMENTS

The author would like to acknowledge valuable discussions with C. Hanke and F. Kappeler.

REFERENCES

/1/ Henry, C.H., Petroff, P.M., Logan, R.A., and Merritt, F.R., J.Appl.Phys. 50 (1979) 3721.

/2/ Yuasa, T. et al., Appl.Phys.Lett. 32 (1978) 119.

/3/ Dobson, C.D. and Keeble, F.S., Symposium on GaAs, London, 1967, pp.68-71.

/4/ Kappeler, F., Mettler, K. and Zschauer, K.-H., IEE Proc. 129 (1982) 256.

/5/ Ueno, M., IEEE J.Quantum Electron. QE-17 (1981) 2113.

/6/ Okajima, M., Watanabe, Y., Nagasaka, H. and Motegi, N. Japan.J.Appl.Phys. 22 (1983) supplem., 329.

/7/ Ungar, J., Bar-Chaim, N. and Ury, I., Electron.Lett. 22 (1986) 279.

/8/ Laidig, W.D., Holonyak, N., Camras, M.D., Hess, D., Coleman, J.J., Dapkus, P.D. and Bardeen, J., Appl.Phys.Lett. 38 (1981) 776.

/9/ Meehan, K., Holonyak, N., Brown, J.M., Nixon, M.A., Gavrilovic, P. and Burnham, R.D., Appl.Phys.Lett. 45 (1984) 549.

/10/ Welch, D.F., Cross, P.S., Thornton, R.L., Burnham, R.D. and Paoli, T.L., Technical Digest IOOC, 1987, paper ME3.

/11/ Shibutani, T., Kume, M., Hamada, K., Shimizu, H., Itoh, K., Kano, G. and Teramoto, I., Technical Digest 10th IEEE Internat. Semiconductor Laser Conf., Kanazawa, 1986, paper K-4.

/12/ Engelmann, R.W.H., Kerps, D., Hom, G. and Ackley, D.E., Technical Digest IEEE Int. Electron. Devices Meeting, San Francisco, 1982, p. 350.

/13/ Kressel, H., Lockwood, H.F. and Hawrylo, F.Z., Appl.Phys.Lett. 18 (1971) 43.

/14/ Hamada, K., Wada, M., Shimizu, H., Kume, M., Susa, F., Shibutani, T., Yoshiawa, N., Itoh, K., Kano, G. and Teramoto, I., IEEE J.Quantum Electron. QE-21 (1985) 623.

/15/ Chinone, N., Nakashima, H., Ikushima, I. and Ito, R., Appl.Opt. 17 (1978) 311.

/16/ Panish, M.B., Casey, H.C., Jr., Sumski, S. and Foy, P.W., Appl.Phys.Lett. 22 (1973) 590.

/17/ Thompson, G.H.B. and Kirkby, P.A., IEEE J.Quantum Electron. QE-9 (1973) 311.

/18/ Tsang, W.T., Appl.Phys.Lett. 40 (1982) 217.

/19/ Garrett, B. and Glew, R.W., Electron. Lett. 23 (1987) 371.

/20/ Tsang, W.T., Electron.Lett. 16 (1980) 939.

/21/ Larsson, A., Salzman, J., Mittelstein, M. and Yariv, A., J.Appl.Phys. 60 (1986) 66.

/22/ Yagi, K., Yamauchi, H. and Niina, T., Technical Digest 10th IEEE Internat. Semiconductor Laser Conf., Kanazawa, 1986, paper K-6.

/23/ Larsson, A., Mittelstein, M., Arakawa, Y. and Yariv, A., Electron.Lett. 22 (1986) 79.

/24/ Welch, D.F., Cross, P.S., Scifres, D.R., Streifer, W. and Burnham,R.D., Appl.Phys.Lett. 50 (1987) 233.

/25/ Nakatsuta, S., Ono, Y. and Kajimura, T., Japan.J.Appl.Phys. 25 (1986) L 498.

/26/ Sakamoto, M. and Kato, Y., Appl.Phys. Lett. 50 (1987) 869.

/27/ Laser Focus/Electro-Optics 23 (1987) 8.

/28/ Laser Focus/Electro-Optics 23 (1987) 24.

/29/ Botez, D. and Ackley, D.E., IEEE Circuits and Device Magazine, Jan. 1986, pp. 8-17.

/30/ Kappeler, F., Siemens Forsch.- u. Entwickl. Ber. 14 (1985) 289.

/31/ Kappeler, F., Westermeier, H., Gessner R., and Druminski, M., Technical Digest 9th IEEE Internat. Semiconductor Laser Conf., Rio de Janeiro, 1984, paper G-3.

/32/ Scifres, D.R., Lindström, C., Burnham, R.D., Streifer, W. and Paoli, T.L., Electron.Lett. 19 (1983) 169.

/33/ Thornton, R.L., Welch, D.F., Burnham, R.D., Paoli, T.L. and Cross, P.S., Appl. Phys. Lett. 49 (1986) 1572.

/34/ Welch, D.F., Scifres, D., Cross, P. and Kung, H., Appl. Phys.Lett. 47 (1985) 1134.

/35/ Harnagel, G.L., Scifres, D.R., Kung, H.H., Welch, D.F., Cross, P.S. and Burnham, R.D., Electron.Lett. 22 (1986) 605.

/36/ Harnagel, G.L., Cross, P.S., Scifres, D.R., Welch, D.F., Lennon, C.R. and Worland, D.P., Appl.Phys.Lett. 49 (1986) 1418.

/37/ Riesinger, A.R., McDonald, P.A., Shealy, J.R., Jochym, E.P., Worth, F., Gilman, J.M., Sprague, J.W. and Zory, P.S., Technical Digest Conf. on Lasers and Electro-Optics, 1987, Baltimore, paper WC 3.

/38/ Lasers & Optronics, late news, May 1987, p.8

/39/ Laser Focus/Electro-Optics Technology 23, April 1987, pp. 74-78.

/40/ Kobayeshi, K. and Mito, I., IEEE J. Lightwave Technol. LT-3 (1985) 1202.

/41/ Oshiba, S., Horikawa, H., Matoba, A., Kawahara, M. and Kawai, Y., Technical Digest 10th IEEE Internat. Semiconductor Laser Conf., Kanazawa, 1986, paper K-1

/42/ Razeghi, M., Blondeau, R., Krakowski, M., de Cremoux, B. and Duchemin, J.P., Appl.Phys.Lett. 50 (1987) 230.

/43/ Figueroa, L., Laser Focus/Electro-Optics 23, April 1987, pp. 18-22.

/44/ Botez, D., IEEE Spectrum 22 (1985) 43.

/45/ Botez, D., Laser Focus/Electro-Optics 23 (1987) 68.

Session A3.2

Thin MOS Bielectrics

Chairman: P. Balk

Wednesday, September 16, 1987

COMPARISON OF METHODS CHARACTERIZING TIME DEPENDENT DIELECTRIC BREAKDOWN IN THIN OXIDE AND OXIDE-NITRIDE-OXIDE LAYERS

P. HIERGEIST, M. KERBER, R. BAUNACH and A. SPITZER

Siemens AG, Corporate Research and Development
Otto-Hahn-Ring 6, D-8000 München 83, FRG

The results of constant voltage stress and constant injection current techniques are discussed concerning dielectric lifetimes and failure modes of a thermal oxide layer and a ONO-layer. The constant voltage stress shows that the ONO-layer has a prolonged lifetime and a lower amount of early failures compared to a single oxide layer, even though the charge to breakdown of the ONO-layer is smaller than that of the thermal oxide. From constant current stress experiments lifetimes for different dielectrics e.g. in a DRAM application can only be inferred in the case of similar electric fields.

1. INTRODUCTION

The development of high density DRAM's requires highly reliable thin dielectric films having low defect density and long lifetimes. Recently new stacked layers e.g. oxide/nitride or oxide/nitride/oxide (ONO) are produced to replace thermal oxide in storage gates of DRAM's [1,2]. The purpose of this work is to discuss the results of constant voltage stress and constant injection current techniques concerning dielectric lifetimes and failure modes of a thermal oxide layer and a ONO-layer. The constant injection current technique is a convenient tool to determine the statistical distribution of the charge to breakdown Q_{bd}, which is used to characterize the failure mode and to compare the quality of thermal oxide layers [3]. However, this characterization method applied to stacked layers containing thermal oxide can lead to errors with regard to dielectric lifetime.

2. PREPARATION OF SAMPLES

Samples for this study were fabricated on (100) orientated p-type silicon wafers. The 20 nm gate oxide was grown in dry oxygen (10% HCl) at 900°C. In the case of the ONO-sample a shallow arsenic implantation was added to the p-type substrate. After thermal oxidation to grow the 7 nm thin bottom oxide layer, LPCVD nitride was deposited and subjected to a wet oxidation. This process results in an 8 nm thin nitride and a 3.5 nm thin oxide layer on top. The oxide equivalent thickness of the ONO-layer is therefore 14.4 nm. In both cases the active areas were defined by a LOCOS process. The gate electrode is a polysilicon layer which was phosphorous doped from the gas phase and structured by photolithography and etching. Capacitors with areas from 10^{-4} mm^2 to 1 mm^2 were available. The samples for the long term high temperature constant voltage stress were bonded in ceramic packages.

3. EXPERIMENTAL RESULTS

In Fig. 1 the cumulative failure of the capacitors with a 20 nm thin thermal oxide is plotted versus the logarithm of the time to breakdown under constant voltage stress for a capacitor area of 1 mm^2 at T=150°C. The distributions show for electric fields above 6 MV/cm two branches with different slopes. At the fieldstrength of 6 MV/cm the steep branch is not found during the time period of 3000 hours. In Fig. 2 the cumulative failure of ONO-capacitors with a 14.4 nm oxide equivalent thickness stressed at the same conditions as the thermal oxide is plotted. The cumulative failure steeply begins to rise after a certain period of time. At a fieldstrength of 10 MV/cm and one second stress duration 34% of the thermal oxide capacitors failed compared to 1% of ONO-capacitors. In Fig. 3 the Q_{bd}-values of 0.01 mm^2 area capacitors of both dielectrics stressed at a current density of 1 mA/cm^2 and T=150°C are plotted in a Weibull diagram. The mean Q_{bd}-value of the ONO-layer is a factor of 100 smaller than that of the thermal oxide. The Q_{bd}-distributions of ONO-capacitors of different areas, measured at T=25°C are shown in Fig.4. For each area the data fit straight lines in a Weibull diagram. Additionally lines with a spacing equal to the logarithm of the ratio of the areas are given.

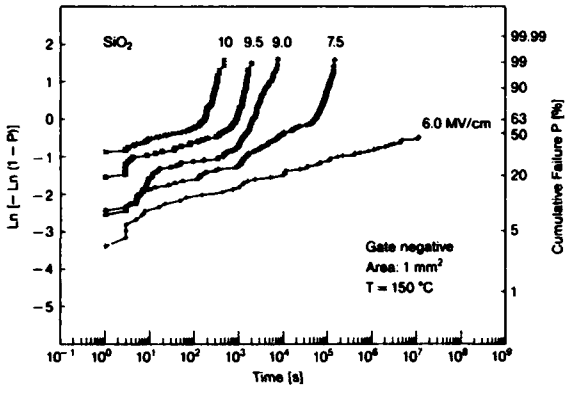


FIGURE 1

Cumulative failure of a thermal oxide (20 nm) at 150°C and different fieldstrengths.

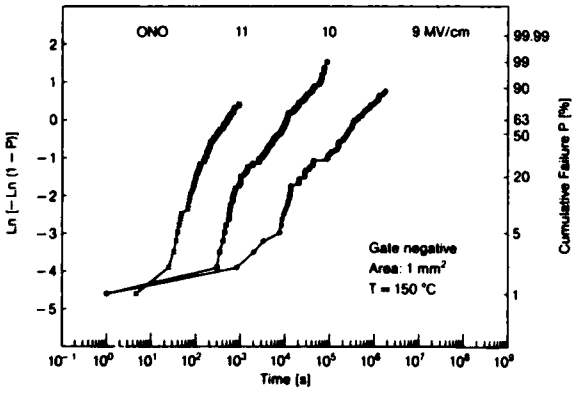


FIGURE 2

Cumulative failure of the ONO-layer (14.4 nm oxide equivalent) at 150°C and different fieldstrengths.

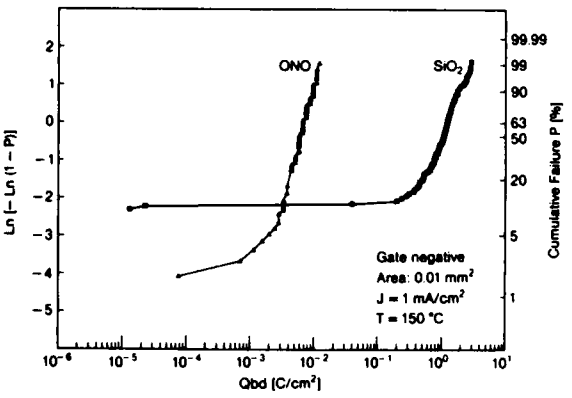


FIGURE 3

Cumulative failure at constant current injection in a thermal oxide (20 nm) and a ONO-layer(14.4 nm oxide equivalent).

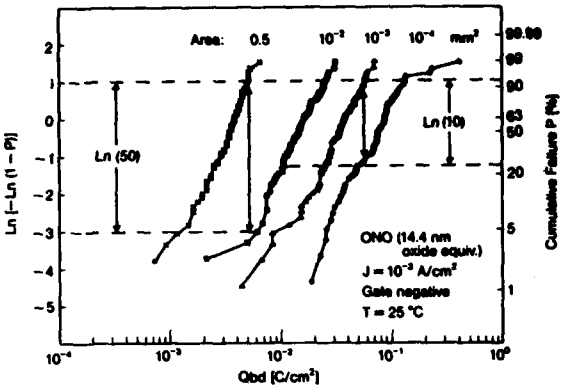


FIGURE 4

Cumulative failure of the ONO-layer (14.4 nm oxide equivalent) at constant current injection. The shift of the failure distribution is proportional to the ratios of areas.

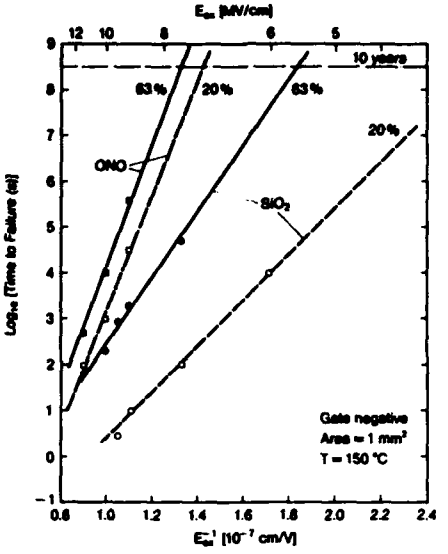


FIGURE 5

Time to failure versus 1/E deduced from the constant voltage stress data of Fig. 2 and 3 for a ONO-layer (14.4 nm oxide equivalent) and a thermal oxide (20 nm).

4. DISCUSSION

The constant field stress shows that the lifetime of the ONO-layer is orders of magnitude higher and the amount of early failures is lower compared to the single oxide layer, even though the Q_{bd} of the ONO-layer is orders of magnitude smaller than that of a thermal oxide with the same electrode area. In Fig. 5 the logarithm of the time to failure deduced from Fig. 1 and Fig. 2 is plotted versus the reciprocal field-strength. In the case of the ONO-layer the straight lines are nearly parallel. This is the consequence of the equal slopes of the failure distributions for different field-strengths in Fig. 2. For the thermal oxide the straigth lines for different failure percentages are not parallel. The lines are inclined against each other since the failure distributions in Fig. 1 are composed of portions with different slopes.

The prolonged lifetimes of the ONO-layer are due to the increased electron trapping compared to the oxide. This reduces the leakage current through the insulator [4]. The trapped electrons in the ONO-dielectric cause an electric field that enhances the potential drop across the oxide layer near the anode [5]. Since the electric fields across both oxide layers are not identical, the anode oxide is preferentially degraded during electrical stress. If the breakdown of the whole triplelayer is initiated by the wearout of the anode oxide the time to breakdown should follow a relationship

$\log t_{bd} = \text{const} \cdot 1/E$

for the field dependence of the oxide breakdown according to the model of the I.C. Chen et al [6]. The

experimental data for the ONO-dielectric agree well with the model confirming the assumptions with respect to the degradation mechanism. Furtheron the results of the constant current stress can be understood on the same basis. During current injection in a ONO-layer it is necessary to apply a higher electric field than for an equivalent oxide layer. This difference is due to the interface properties within the triplelayer dielectric [5]. Since the highest potential drop is across the oxide layer near the anode, the Q_{bd}-measurements are testing an oxide wearout at very high electric fields (above 15 MV/cm) with the consequence of smaller Q_{bd}-values compared to a single oxide layer.

The Q_{bd}-values of a ONO-layer depend on the gate area (Fig. 4). The cumulative failure is Weibull-distributed with regard to the break-down charge. The cumulative failure shifts along the vertical axes of the Weibull diagram according to Log (A_1/A_2) (A: capacitor area). Therefore the breakdown phenomena is due to a mechanism related to extrem value statistics.

5. CONCLUSION.

The ONO-layer has reduced early failures and a lifetime, which is orders of magnitude higher compared to the single oxide layer . The cumulative failure of the ONO-layer shows a Weibull-type of failure distribution with a small deviation of Q_{bd}. The field acceleration behaviour of the ONO-layer agrees well with the model of I.C. Chen et al. for thin oxides. From constant current stress experiments a comparison of device lifetimes e.g. in a DRAM application can only be extracted in the case of similar electric fields.

ACKNOWLEDGEMENT

This report is based on a project which has been supported by the Minister of Research and Technology of the Federal Republic of Germany under the support-no. NT 2696. For the contents the authors alone are responsible.

REFERENCES

[1] D.S. Yaney, J.C. Desko, M.J. Kelly, L.T. Lancaster, A.M-R Lin, A.S. Manocha, T.K. Mc Guire, F.R. Pfeiffer, and H.C. Kirsch, IEDM Tech. Digest (1985) 698

[2] T. Watanabe, A. Menjoh, M. Ishikawa, and J. Kumagei, IEDM Tech. Digest (1984) 173

[3] J. van der Schoot and D.R. Wolters, Current induced dielectric breakdown, in: J.F. Verwey and D.R. Wolters (eds.), Insulating films on semiconductors (North Holland, Amsterdam, 1983) pp. 270-273

[4] D.J. Di Maria, D.R. Young and D.W. Ormond, Appl. Phys. Lett, Vol 31

[5] R. Baunach and A. Spitzer, to be published in Appl. Surf. Sci.

[6] I.C. Chen and Ch. Hu, IEEE EDL-8 (1987) 140

A3.2.2

COMPARISON OF THE INTERFACIAL STRESS RESISTANCE IN RAPID THERMALLY PROCESSED THIN DIELECTRICS

R B Calligaro, P J Rosser* and P B Moynagh*

GEC Research Limited, Hirst Research Centre, East Lane, Wembley, Middlesex, HA9 7PP United Kingdom

*STL Technology Ltd., London Road, Harlow, Essex, United Kingdom

The dielectric-silicon interface stress resistance of conventional furnace oxides and those formed using rapid thermal oxidation (RTO) with and without rapid thermal nitridation (RTN) are compared. The stress resistance is deduced from the change in the fast interface trap density and flatband voltage under high field constant current stress conditions. In the thinner dielectrics (10-13 nm), the interfacial stress resistance of the RTN dielectrics at ~9 MV/cm is a factor of 2 greater than in the thermal oxides. For the thicker RTN layers (34-40 nm) the stress resistance was ~25% less than the thermal oxides due to the reduced interfacial nitrogen concentration. However in the thinner dielectrics, the higher nitrogen concentration in the RTN layers leads to a factor of ~8 greater shift in flatband voltage compared to the thermal oxides.

1. INTRODUCTION

Considerable attention has been focussed on alternative gate dielectrics to silicon dioxide for submicron MOS devices, particularly with regard to their hot electron injection resistance. The most promising of these dielectrics is nitrided thermal oxide because of its compatibility with conventional processing and resistance to trap formation. However, the high energy requirements to form these layers have precluded their general use. This disadvantage can be minimized by using rapid thermal processing to reduce the thermal budget yet maintain good dielectric quality [1]. Previous work [2] has found that nitrided oxides produced by conventional furnace techniques exhibit a low trap generation rate during stress. However, the behaviour of the interface during stress has not been examined for dielectrics formed by rapid thermal processing. In this work, the stress resistance of the Si-dielectric interface is compared for conventional furnace oxides, and those formed by rapid thermal oxidation (RTO) with and without rapid thermal nitridation (RTN).

2. SAMPLE PREPARATION AND MEASUREMENT

Conventional oxides were grown in dry oxygen in a resistance heated furnace and rapid thermal processing was carried out in a Heatpulse 410T machine. Substrates used were <100> n-type of resistivity 0.1-4 Ω-cm. Dielectric thicknesses in both cases were restricted to 10-13 nm and 34-40 nm. Conventional oxides were grown at 950°C and RTO samples at 1100°C. Nitridation was carried out on RTO samples at temperatures of 1200°C for 60sec and 120sec in an ammonia ambient. Polycrystalline silicon of 0.4 μm thickness was deposited at 620°C and doped with $POCl_3$ at 950°C for 20 min on all the wafers. Aluminium was deposited by magnetron sputtering at 200°C and 200 μm diameter capacitors were photo-engraved into the underlying layers. At the metal sinter stage the ambient used was forming gas (10% H_2 in N_2) for 30 min at 435°C.

The dielectric layers were stressed using constant current, high field injection techniques. Stress conditions were selected to inject a constant charge (10 mC/cm^2) at current densities of 200 nA/cm^2 to 1mA/cm^2 or a constant current of 20 $\mu A/\mu m^2$ up to an injected

charge of 100mC/cm^2.

Immediately after the stress period, 100 kHz high frequency and 50 mV/sec ramp rate quasistatic capacitance-voltage curves were obtained and analysed to give the fast interface trap density as a function of energy in the band gap. An example of the curves obtained for the 10-13 nm thick samples is shown in Figure 1a. The calculated interface

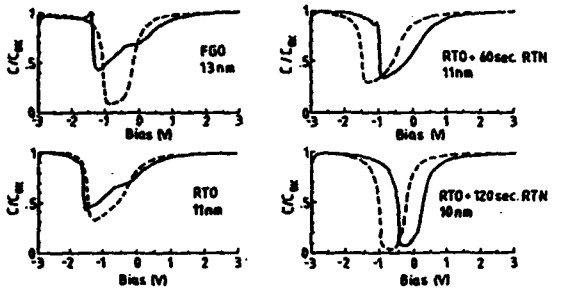


FIGURE 1a. PRE AND POST STRESS C-V CHARACTERISTICS OF 10-13nm THICK DIELECTRICS. (pre dashed, post solid line)

trap density distribution is shown in Figure 1b

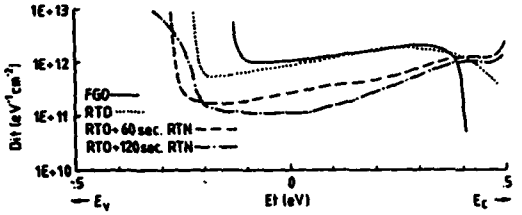


FIGURE 1b. FAST INTERFACE TRAP INTENSITY OF DIELECTRICS IN FIGURE 1a.

and the distribution for the thick oxide samples (34-40 nm) is shown in Figure 1c.

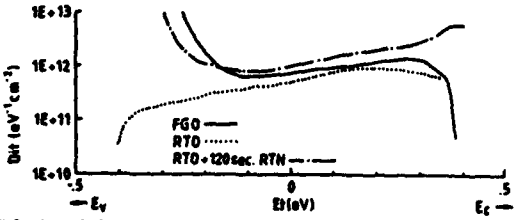


FIGURE 1c. FAST INTERFACE TRAP DENSITY OF 34-40 nm DIELECTRICS

3. RESULTS AND DISCUSSION

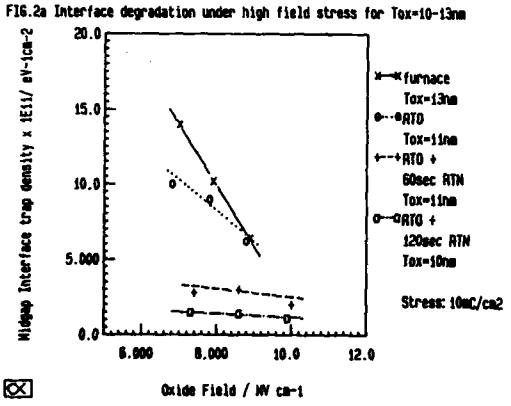
Prior to the high field constant current stress a selection of capacitors of each dielectric type and thickness were measured with regard to fixed oxide charge, fast interface trap density at midgap, dielectric -silicon barrier height and breakdown fields. The fixed oxide charge in the 60sec and 120 sec RTN dielectrics was in the range 2.1 -3.4 10^{11}cm^{-2} compared to <5 x 10^{10}cm^{-2} for the furnace and RTO oxides. The fast interface trap density at midgap for all the dielectrics was low and within the resolution limit for the particular substrate doping level [3]. From the I-V measurements the barrier height in the 10-13 nm thick RTN dielectrics was calculated to be 0.3eV less than the thermal oxides (2.9 eV). However, for the 34-40 nm thick RTN dielectric, the barrier height was equivalent to the thermal oxide samples indicating a low concentration of nitrogen at the interface. The breakdown field of all the thin 10-13 nm dielectrics was 13 MV/cm and this was reduced by ~15% to 11 MV/cm in the thicker samples.

After stressing at high field with constant current, the original C-V characteristics are distorted by the formation of fast interface traps and negative charge in the dielectric. Injection of charge to 10 mC/cm^2 at 20 μA/cm^2 increases the minimum capacitance and the width (at C/Cox = 0.9) of the C-V characteristics as shown in Figure 1a. The positive shift in the flatband voltage is greater in the RTN dielectrics than in the thermal oxides indicating a higher density of trapping centres within the nitrided dielectric. This is discussed in more detail in subsequent sections. However the degradation of the dielectric-silicon interface is least in the RTN samples as seen by the small change in the width and minimum of the C-V curves. The RTN samples are resistant to the formation of discrete interface trapping centres at 0.3 eV above midgap as can be seen by the absence of the ledge in the C-V curves in weak depletion. Doubling the nitridation time to 120 sec, additionally removes the very slow interface traps which lead to a non-equilibrium peak response in the C-V curves of the remaining

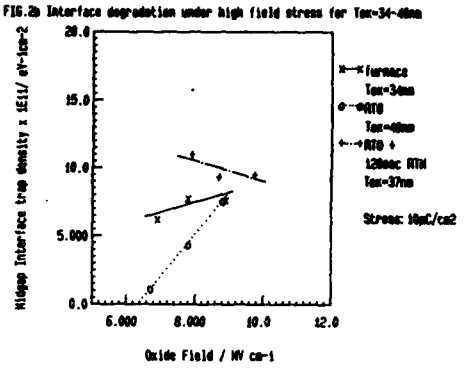
samples. Extraction of the fast interface trap density, D_{it} using the high-low or ideal capacitance method [3] as shown in Figure 1b confirms the excellent stress resistance of the thin nitrided samples. However, for the thicker dielectric samples (34-40 nm) the 120sec nitridation is not sufficiently long to provide a comparable stress resistance to the thinner layers (see Figure 1c).

The above measurements were extended to stress current densities of 1 mA/cm^2 and 200 nA/cm^2 for the same injected charge density of 10 mC/cm^2. Therefore the interface degradation measured at midgap can be obtained as a function of the oxide field and these results are plotted in Figure 2 for the 10-13 nm layer (Figure 2a) and the 34-40 nm dielectrics (Figure 2b). A low value of injected charge was used to prevent significant charging of traps within the dielectric and thus reduce the distortion of the electric field.

In Figure 2a the stress resistance of the

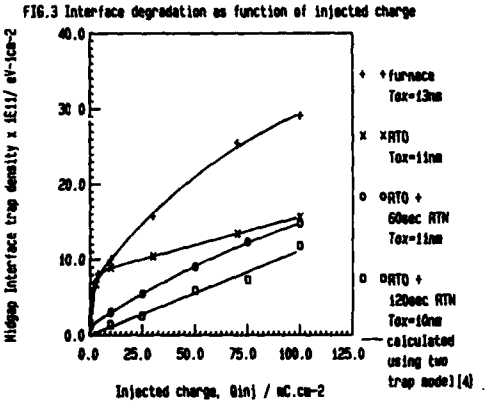


120sec RTN dielectric is a factor of 2 higher than the 60 sec RTN equivalent. In both cases the interface degradation exhibits a significantly smaller change with increasing field than the thermal oxides and the RTN samples show superior stress resistance in comparison to the thermal oxides. At low fields, the stress resistance of the RTO



samples is greater than the furnace grown oxide. This is also true for thicker layers as shown in Figure 2b. However, in this case the stress resistance of the nitrided layer, although changing significantly less than the thermal oxides with increasing field, shows a ~25% increase in degradation at ~9 MV/cm in comparison to the thermal oxides. This may be due to either a difference in interfacial stress or hole barrier height due to the smaller nitrogen concentration (~4%) at the dielectric-silicon interface compared to the thinner nitrided samples (~8%) as measured by Auger emission spectroscopy.

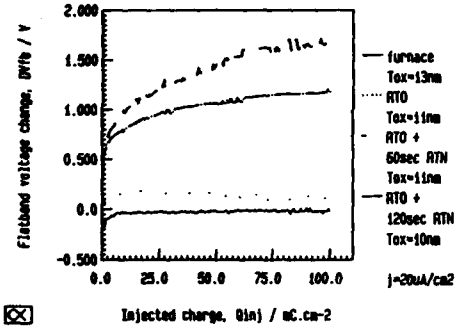
For the thin dielectric layers, the interface degradation and the flatband voltage change as a function of injected charge at a fixed current density were also measured. These results are displayed in Figures 3 and 4. In



confirmation of the results for constant injected charge, the minimum interface degradation is obtained for the 120 sec RTN dielectric. Additionally, the rate of change of the interface degradation is significantly less in the RTN dielectrics. Therefore, in MOSFETs the gain and subthreshold slope degradation of the turn on characteristics is expected to be minimised if RTN gate dielectrics replace thermal oxides.

The change in the flatband voltage, V_{fb}, with injected charge is related to the density of neutral traps in the dielectric and a positive shift in the V_{fb} as shown in Figure 4 indicates negative charge storage. These curves give a relative measure of the expected change in the threshold voltage of a MOSFET under stress. Therefore, as can be seen from Figure 4, the

FIG.4 Flatband voltage change as function of injected charge



RTN dielectrics contain a higher density of neutral traps than the thermal oxides. This results in flatband voltage shifts which are a factor of ~8 greater than in the thermal oxides. Consequently, MOSFETs fabricated with RTN dielectrics, although gaining resistance to subthreshold slope and gain degradation, are expected to be less resistant to threshold voltage shift than equivalent thermal oxides. These results are in agreement with previous work [5] where low pressure nitridation of oxide films was studied. It was found that appropriate re-oxidation of the nitrided oxide removed electron traps, thus reducing the Vt shift degradation, yet maintained the interface integrity under stress. Similar techniques, applied to nitrided RTO films, would also be expected to achieve comparable results.

3. CONCLUSIONS

The dielectric-silicon interface stress resistance of conventional furnace oxides and those formed using rapid thermal oxidation with and without rapid thermal nitridation has been compared. In the thinner dielectrics of thickness 10-13 nm the interfacial stress resistance of the RTN dielectrics is a factor of 2 greater than the thermal oxides at ~9 MV/cm. However, the flatband voltage change of the RTN dielectrics was found to be a factor of ~8 greater than in the thermally oxidised samples.

4. ACKNOWLEDGEMENT

This work has been partly supported by the Alvey Directorate.

5. REFERENCES

1. Nulman J, Krusius J P and Rathbun L, IEEE Proc. of IEDM, 1984, p169-172
2. Ekstedt, T W, Wong S S, Strausser Y E, Amano J, Kwan S J and Gronolds H R, Proc INFOS 83, Elsevier Science Publishers BV. pp189-193
3. Nicollian E H and Brews J R, MOS Physics and Technology, New York J Wiley and Sons 1982
4. Fischetti M V, J. Appl. Phys. 56, 1984, pp 575-577
5. Jayaraman R, Yang W, and Sodini C G, Proc. IEEE IEDM 1986, pp668-671

ELECTRICAL CHARACTERISTICS AND RELIABILITY OF THIN OXIDE-NITRIDE-OXIDE STACKED FILMS

L. Do Thanh and P. Balk

Institute of Semiconductor Electronics
Aachen Technical University, D-5100 Aachen, FRG

Electronic conduction, charge trapping and dielectric breakdown have been studied on thin stacked layers of SiO_2-Si_3N_4-SiO_2 with Al or poly-Si gate on Si. These structures appear to combine the attractive properties of SiO_2 and Si_3N_4 single layers as long as they are operated at fields below 5 $MVcm^{-1}$.

1. INTRODUCTION

Storage capacitors for dynamic RAMs of 4 Mbit and beyond require very thin insulator films to obtain sufficiently large capacitance at small cell size /1/. The use of single SiO_2 or Si_3N_4 films would lead to measurable conduction at the thicknesses and fields needed for this application /1,2/. Stacked SiO_2-Si_3N_4-SiO_2 triple layers appear to be a promising alternative /3/ since they would combine the advantages of Si_3N_4 layers (large breakdown field strength, excellent wear-out behavior /4/) with those of SiO_2 layers (high barriers for carrier injection and low trap density at Si-SiO_2 interface). Results on the properties of such stacks are the topic of the present contribution.

2. EXPERIMENTAL

MIS capacitors were prepared on p- and n-type 0.1 cm (100) Si wafers. First a thin (6 to 9 nm) SiO_2 layer was thermally grown (O_2, 900°C) or deposited (N_2O-SiH_4, 250°C, PE-CVD). A Si_3N_4 film was deposited next (10 to 11 nm, SiH_2Cl_2-NH_3, 825°C, LP-CVD). The top oxide layer (7 nm) was obtained by thermal oxidation (wet O_2, 900°C) or deposition (PE-CVD). Gate electrodes were Al or P-doped poly-Si. Finally, all samples were annealed at 400°C in N_2 for 20 min.

3. RESULTS AND DISCUSSION

Contrary to high frequency CV curves of MIS capacitors with single Si_3N_4 films those of structures with stacked dielectrics do not show hysteresis. This indicates that low field charge injection does not occur. The fixed charge is negligible ($< 2x10^{10} cm^{-2}$) and the interface trap density in the low $10^{10} cm^{-2}eV^{-1}$ region. For all samples and both polarities fields > 5 $MVcm^{-1}$ are necessary to reach an injection current level > $10^{-8} Acm^{-2}$. At a given field the current is several orders of magnitude lower than in single Si_3N_4 films. Like for SiO_2-Si_3N_4 double layers with oxide films over 5 nm thick /5/ the dominant carriers are probably electrons. In this study the fields were calculated by subtracting the flatband voltage from the gate voltage and dividing the result by the equivalent SiO_2 thickness, i.e. the SiO_2 thickness which yields the same capacitance as the stack. The current at positive gate voltage is always the larger one; it increases more strongly with field than the current at negative gate. This suggests stronger electron trapping in the latter case. It may be expected that the potential well nature of the triple layer structure plays an important role in the trapping.

Breakdown field strengths for stacked insulator structures with Al gate measured by stressing the samples in accumulation using a fast ramp are shown in fig. 1. The mean destructive breakdown field is > 10 $MVcm^{-1}$. Samples two PE-CVD SiO_2 layers yield a distinctly broader distributions than those with one or both SiO_2 films prepared by thermal oxidation. Tight distributions with somewhat lower mean breakdown fields were obtained with poly-Si gate samples using thermal oxides (fig. 2).

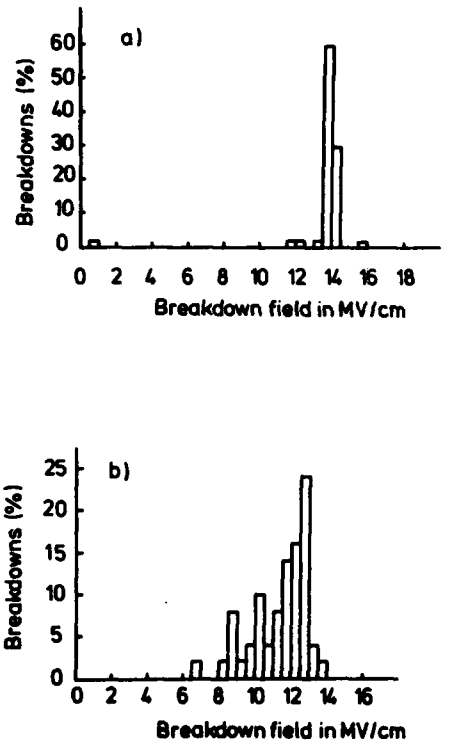


FIGURE 1

Breakdown histograms of Al gate structures with stacked insulators and p-Si substrate (Vg negative); samples:
a) Al-therm. SiO_2 (6.3 nm) - Si_3N_4 (7.4 nm) - therm. SiO_2 (6.0 nm) - p-Si
b) Al-PECVD SiO_2 (6.9 nm) - Si_3N_4 (10 nm) - PECVD SiO_2 (6.7 nm) - p-Si.

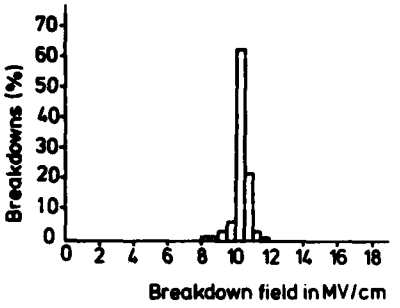


FIGURE 2

Breakdown histograms of poly-Si gate structures with stacked insulators on p-Si substrate (V_g negative); sample: poly-Si gate - therm. SiO_2 (6.3 nm) - Si_3N_4 (6.3 nm) - therm. SiO_2 (8 nm) - p- Si.

Wear-out properties of MIS systems with stacked insulators were studied by stressing at constant current. Charge trapping during this process was monitored by measuring the shifts of the gate voltage and the flatband voltage. The data in fig. 3 show that for both polarities the gate voltage must be increased to keep the current constant, which indicates the capture of negative charge. However, the change in gate voltage at negative gate is larger than at positive gate. Similar results are obtained for the flatband voltage. These results appear to indicate again that more negative charge is captured at negative than at positive gate polarity. The centroid of charge must be located closer to the substrate than to the gate because the flatband voltage shift exceeds that of the gate voltage.

For larger injected charge densities at negative gate polarity the flatband voltage shift turns around, indicating the build-up of positive charge. This effect has also been reported by other authors /6/. At these conditions the density of interface traps was observed to increase. A larger increase in the interface trap density was found at positive gate polarity, even though the turn around

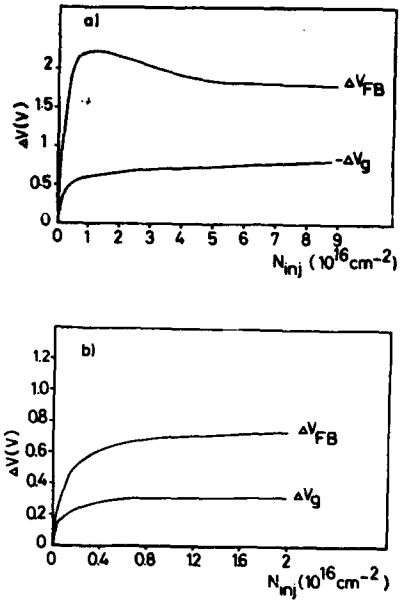


FIGURE 3

Shift of flatband (ΔV_{FB}) and gate voltage (ΔV_g) during constant current stressing of poly-Si gate structures with a) negative ($J = 10^{-5}A/cm^2$) and b) positive ($J = 10^{-4}A/cm^2$) gate polarity; samples: poly-Si - therm. SiO_2 (6.3 nm) - Si_3N_4 (6.3 nm) - therm. SiO_2 (8 nm) - p-Si (a) or n-Si (b). Note that in (a) ΔV_g is negative.

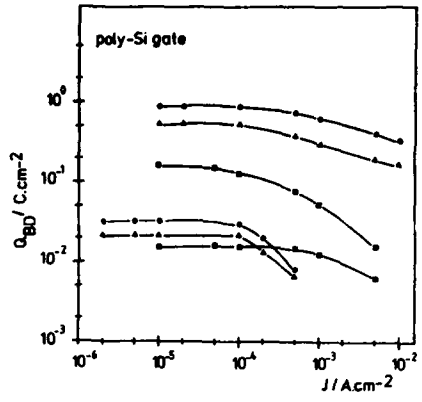


FIGURE 4

Density of injected charge to breakdown vs. density of injected current; samples:
(o, •): poly-Si - therm. SiO_2 (6.3 nm) - Si_3N_4 (6.3 nm) - therm. SiO_2 (8 nm) - n-Si
(Δ, ▲): poly-Si - therm. SiO_2 (2.9 nm) - Si_3N_4 (8.2 nm) - therm. SiO_2 (8 nm) - n-Si
(□, ■): poly-Si - therm. SiO_2 (6.3 nm) - Si_3N_4 (6.3 nm) - therm. SiO_2 (8 nm) - p-Si
Open symbols: V_g positive; filled symbols: V_g negative

effect was not observed. It is likely that the turn around at negative gate is caused by secondary injection of holes from the Si substrate due to field enhancement by electrons trapped in the nitride.

We have shown in a previous study /4/ that the density of injected charge (Q_{BD}) leading to breakdown of LP-CVD Si_3N_4 films (approx. 10^4C cm^{-2}) is larger by several orders of magnitude than that for thermal SiO_2 films (approx. 10 C cm^{-2}). For stacked dielectrics Q_{BD} only attains a value of approx. 1 C cm^{-2} in some cases, but is often considerably less. Wear-out data on poly-Si gate samples with stacked insulators (fig. 4) show that Q_{BD} is systematically larger for positive than for negative gate voltage. We suggest that the low Q_{BD} values in the latter case may be related to the above mentioned secondary injection of holes. At high currents (high fields) the value of Q_{BD} decreases. The larger fields required to obtain a given current density, as discussed earlier, are not sufficient to explain the small Q_{BD} at negative gate since even at the same magnitude of electric field Q_{BD} is still considerably smaller for this polarity.

For Al gate samples similar trends were observed. Again, the highest Q_{BD} values were approx. 1 C cm^{-2}, but those for negative gate bias were one order of magnitude lower than in the poly-Si gate case. The dependence on the current level is less than for poly-Si gates. The use of two PE-CVD SiO_2 films reduces Q_{BD} also for positive stress voltage to below 10^{-1} C cm^{-2}. It is interesting to note that the use of a very thin top electrode does not affect Q_{BD}, as shown for poly-Si gate samples in fig. 4. Injection of holes from the gate at positive gate polarity should be possible in this case /5/. However, these carriers would not be able to leave the potential well and reach the Si-SiO_2 interface.

4. CONCLUSIONS

Stacked SiO_2-Si_3N_4-SiO_2 insulators exhibit a number of attractive properties for application in storage capacitors. They are less conductive than Si_3N_4 films and show a lower interface trap density. Stacked films have a higher dielectric constant than SiO_2 but a comparably high breakdown field. However, care should be taken to operate these capacitors below 5 $MVcm^{-1}$ in order to avoid charging problems caused by the potential well nature of these structures. The amount of injected charge leading to breakdown is smaller than that for single SiO_2 or Si_3N_4 layers.

References

/1/ Noble, W.P. and Walker, W.W., IEEE Circuits and Devices Magazine, 1, 45 (1985)

/2/ Andrews, J.M., Jackson, B.G. and Polito, W.J., J. Appl. Phys., 56, 1587 (1986)

/3/ Watanabe, T., Menjok, A., Ishikawa, M. and Kumagai, J., IEDM Techn. Digest, 1984, p. 173

/4/ Do Thanh, L. and Balk, P., in Proceedings of the "Symposium on Silicon-Nitride and Silicon Dioxide Thin Insulating Films" E.C.S., Princeton N.J., 1987), to be published

/5/ Maes, H.E. and Heyns, G., in "Insulating Films on Semiconductors", eds. J.F. Verweij and D.R. Wolters, (North-Holland, Amsterdam 1983), p. 215

/6/ Nozaki, S. and Giridhar, R.V., IEEE Electron Dev. Lett., EDL-7, 486 (1986)

TRAPPING PROPERTIES OF VERY THIN NITRIDE/OXIDE GATE INSULATORS

J.Y-C Sun, M.Arienzo, L.Dori†, and K.Stein

IBM Thomas J.Watson Research Center,
P.O.Box 218, Yorktown Heights, N.Y. 10598, USA

The trapping properties of very thin nitride/oxide (10-14nm equivalent SiO_2) composite gate insulators and their dependences on gate materials and process conditions are reported. Electron trapping and flatband voltage turn-around effects are more pronounced in these films than in thermal SiO_2. They both appear to be dominated by water-related species in the bottom oxide layer when the top nitride layer is thin, similar to the case of thermal SiO_2 only. For VLSI CMOS applications, trapping and instabilities in the nitride/oxide gate insulator can be minimized by (i) reducing the thickness of the top nitride layer, (ii) using polysilicon gates with proper work functions, and (iii) using appropriate high-temperature dehydration steps after polysilicon gate deposition.

1. INTRODUCTION

High integrity nitride/oxide composite gate dielectric structures have received high technological interest as potential substitutes for SiO_2 in VLSI-IGFETs and memory devices such as DRAMs and non-volatile memories [1-3]. Such multilayered gate dielectric structures with a thin nitride film (< 5 nm) on top of 8-10 nm of SiO_2 offer enhanced diffusion resistance, higher process tolerance, tight spread on breakdown field as well as breakdown charge density distributions, and lower drain-hot carrier induced V_T shifts [2,3]. Moreover, the good interfacial properties of silicon dioxide on silicon are preserved.

However, charge trapping in the top nitride layer has been one of the major road blocks to the success of such gate insulators [2]. The goal of this work is to study the trapping properties of such thin gate insulators. New insights into the gate material and processing dependences of charge trapping in thin nitride/oxide composite gate insulators are obtained, which is crucial to the successful application of such insulators in VLSI CMOS.

2. EXPERIMENTAL

The nitride/oxide (NO) composite gate insulator consists of a thin (4 - 7 nm) LPCVD nitride layer on top of a thermally grown thin oxide (8-9 nm) on a <100> Si substrate. The LPCVD nitride was deposited immediately after the oxidation. The nitride deposition rate was reduced by increasing the ratio of ammonia to dichlorosilane to provide good thickness control.

N and P channel IGFETs with either thermal oxide (OX) or the composite insulator (NO) were fabricated without threshold-adjust ion implantation. The background (substrate) doping is 8×10^{15} cm^{-3} for the n-channel IGFET and 8×10^{16} cm^{-3} for the p-channel IGFET, respectively. The polysilicon gate was degenerately doped to n-type for n-channel FETs and p-type for p-channel FETs respectively, by the source/drain ion implantation and drive-in [4].

Nitride thickness was evaluated by ellipsometry, using a double-absorbing-layer program. The measured refractive index was 1.9, which indicates the non-stoichiometry of this very thin nitride film. To characterize the trapping properties of the NO structure, two low-field charge injection techniques have

† Permanent address: CNR-Instituto LAMEL, Via Castagnoli, 1 - I-40126 Bologna, Italy

been used: a) avalanche electron injection performed on MOS capacitors, with the substrate implanted with boron, as reported in [5] to obtain a uniform injection. In this case the polysilicon gate was either As *in situ* doped with no heat treatment after polysilicon deposition, or As ion-implanted and followed by a rapid thermal anneal at 1000°C for 6 sec; and b) refined substrate hot electron (hole) injection technique on IGFETs devices [6].

3. RESULTS

Fig. 1 reports V_{FB} shift as a function of injection time for 4 nm nitride on top of 8 nm SiO_2, and, as a reference, for 10 nm SiO_2 alone. The trapping properties of the composite insulator are significantly different from those of SiO_2 only. In fact, the 10 nm oxide exhibits a small V_{FB} shift with a saturation value ($\simeq$ 0.35 V), in agreement with the t_{ox}^2 dependence of the V_{FB} shift as reported by Young *et al.* [7] for dry oxides down to 20 nm. On the contrary, the NO structure has a higher trapping efficiency. Further, V_{FB} shift as a function of injection time shows a quick turn-around a few seconds after the injection is started.

The amount of electron trapping and V_{FB} turn-around in NO films depends strongly on gate electrode materials as well as processing conditions after the NO gate insulator is grown. Such dependences are shown in Fig. 2, where three different cases are compared: (1) aluminum gate, (2) arsenic *in situ* doped n^+ polysilicon gate without any subsequent high temperature processing, and (3) arsenic ion-implanted polysilicon gate with subsequent rapid thermal anneal (RTA) at 1000°C for 6 seconds in dry nitrogen. Among these three, case 1 (Al-gate) has the most electron trapping and V_{FB} turn-around effect, while case 3 (1000°C RTA) has the least.

To eliminate the turn-around effect (see Fig. 1 and 2) and to study the electron traps alone, the substrate temperature during injection should be increased over 323°K. The results are summarized in Fig. 3, where V_{FB} shifts as a function of injection time at different temperatures are reported. From this plot, information about the centers responsible for the trapping of the positive charge can be obtained by subtracting the curve corresponding to the injection at 293°K from that at 373°K. In agreement with [8], a very small positive charge trap is found. Its cross section (10^{-18} cm^2) is comparable to that of the anomalous positive charge (APC) trapping in SiO_2 [9]. The density of this small trap is high, roughly 10^{12} cm^{-2}, in NO structures with RTA polysilicon gate electrodes (case 3).

In contrast, no turn-around phenomenon was observed on NO structures with a 4 nm top nitride layer during substrate hot electron (SHE) injection in IGFET devices as shown in Fig. 4, where V_T shift is reported as a function of injected charge. It is clear that electron trapping increases with the thickness of the top nitride layer. Further, in Fig. 4, the V_T shift in the case of a MOSFET with a polysilicon gate electrode and 10 nm of SiO_2 is very small and much less than that of the MOS capacitor with an aluminum gate electrode. As far as hole trapping is concerned, under substrate hot hole (SHH) injection the NO structure exhibits a lower V_T shift than the 10 nm oxide.

4. DISCUSSIONS

As seen in Fig.1, there is much more pronounced electron trapping and V_{FB} turn-around in the NO composite insulator than in thermal SiO_2. Undoubtedly, higher electron trapping efficiency is the characteristic of LPCVD nitride films which tend to be not as dense as thermally grown films.

On the other hand, it is well known that electron trapping in thermal SiO_2 is enhanced by the presence of water-related species [10]. The electron capture cross sections of water-related traps are typically in the range of 10^{-17} to 10^{-18} cm^2 [10]. Analyses of the 150°C injection data in Fig. 3 show that the capture cross section of the dominant electron trap in these very thin NO films is roughly 10^{-17} cm^2, in agreement with that of the water-related traps in the thermal SiO_2. Since hydrogen and/or water is naturally present in the LPCVD nitride process ambient, the incorporation of water-related species or hydrogen in the underlying SiO_2 film during the initial period of

thin nitride deposition is very likely. We therefore believe that the increased number of water-related traps in the bottom oxide layer due to nitride deposition is another major contributing factor to the enhanced electron trapping in the composite NO insulator. The reduction of electron trapping with polysilicon gate electrodes and subsequent high temperature anneals in a dry inert ambient (Fig. 2.) is consistent with the dehydration of the bottom thermal oxide. Besides, the high temperature anneal can also densify the nitride and thus reduce the electron trapping in the top nitride layer.

There are two factors contributing to the V_{FB} turn-around effect in the nitride/oxide composite insulator. The first one is hole injection from the gate electrode into the nitride layer via trap-assisted tunneling [3], followed by subsequent trapping of injected holes in the nitride. This effect has been known as the threshold voltage instability under dc gate bias [3,4]. However, the magnitude of the negative V_{FB} shift under dc gate bias [3] is not sufficient to account for the large V_{FB} turn-around effect observed during avalanche electron injection. The second and perhaps the dominant factor in V_{FB} turn-around is the same anomalous positive charge generation that was observed in thermal oxide alone [7,8]. In fact, the dependences of electron trapping and V_{FB} turn-around effects in the NO film on gate materials and processing conditions (Fig. 1 and 2) follow exactly the same trend as those of thermal SiO_2 [11]. The elimination of the turn-around effect during injection at elevated temperatures (150°C, Fig. 3) is also identical to that of thermal SiO_2 [7,8]. These similarities strongly suggest that the V_{FB} turn-around effect in the NO film during avalanche electron injection is most likely dominated by the anomalous positive charge generation in the bottom oxide layer. Although there are many models proposed to explain the origin of the anomalous positive charge, for example, a notable one is donorlike surface states [11], there is little doubt that it is enhanced by water or hydrogen incorporation in the oxide [7,11]

It is clear from Fig. 2 and 4 that electron trapping as well as the V_{FB} turn-around effect in such NO composite insulators can be minimized by reducing the thickness of the top nitride layer and by using polysilicon gates with subsequent high temperature anneals in a dry inert ambient. In this respect, thin top nitride layer not only reduces the trapping in itself but also allows the dehydration process of the bottom oxide layer to proceed during polysilicon deposition and subsequent high temperature anneals.

As far as hole trapping is concerned, net hole trapping or negative V_T shift during substrate hot-hole injection appears to be less in the NO film than in thermal SiO_2 as Fig. 4 indicated. This could be due to the injection and trapping of electrons from the gate electrode in the top nitride layer and the recombination of trapped holes with injected electrons.

5. CONCLUSION

The trapping properties of very thin nitride/oxide composite gate insulators have been established by two low-field charge injection techniques: avalanche electron injection (for the first time), and refined substrate hot-electron (hole) injection. Electron trapping and flatband voltage turn-around effects are more pronounced in these NO films than in thermal SiO_2. Electron trapping in the nitride dominates the flatband or threshold voltage shift when the top nitride layer is thick (>7 nm). In contrast, water-related traps in the bottom oxide become more important when the top nitride layer is thin (4 nm). However, trapping and instabilities in these NO gate insulators can be minimized for VLSI-CMOS applications by (i) reducing the thickness of the top nitride layer, (ii) using polysilicon gates with proper work functions, and (iii) using appropriate high temperature anneals after polysilicon gate deposition.

ACKNOWLEDGMENTS

We thank D. Zicherman, M. Rodriguez, J. Calise, and T. Nguyen for testing support, and the IBM Yorktown Silicon Facility for processing support.

REFERENCES

[1] J.J. Chang, Proc. of IEEE, p. 1039 (1979)

[2] K.K. Young *et al.* , VLSI Tech. Symp., p. 65 (1986)

[3] J. Y.-C. Sun, *et al.* , Proc. Int. Symp. on VLSI Tech. Sys. and Appl., p. 106 (1987)

[4] J. Y.-C. Sun *et al.* , IEDM Tech. Dig., p. 236 (1986)

[5] L. Dori, *et al.* , J. Appl. Phys., v. 61, 1910 (1987)

[6] T.H. Ning, *et al.* , J. Appl. Phys., v. 49, p. 5997 (1978)

[7] D.R. Young, *et al.* , J. Appl. Phys., v. 50, p. 6366 (1979)

[8] B.H. Yun, Appl. Phys. Lett., 27, 256, (1975)

[9] M.V. Fischetti, J. Appl. Phys., 57, 2860 (1985)

[10] D. DiMaria, in: Physics of SiO_2 and its Interfaces, (ed. by S.T. Pantelides, Pergamon Press), p. 160, (1978).

[11] C.T. Sah, *et al.* , J. Appl. Phys., v. 55, p. 1525 (1984)

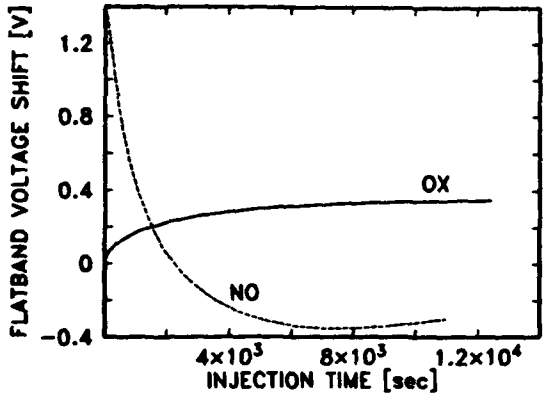


Fig. 1. V_{FB} shift as a function of injection time at room temperature for the 10 nm SiO_2 (OX) and 4 nm nitride + 8 nm oxide (NO). Current density = 1.93×10^{-5} A/cm^2. Gate material = Al.

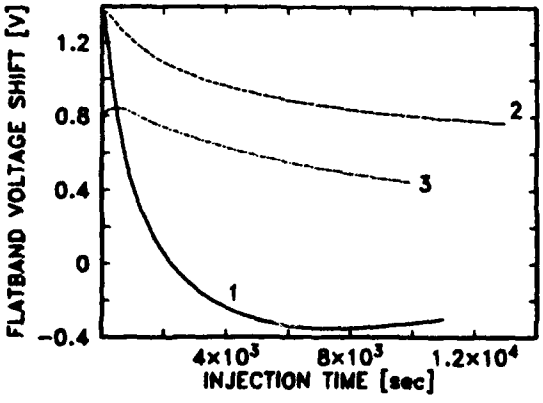


Fig. 2. V_{FB} shift as a function of injection time at room temperature for the 4 nm nitride + 8 nm oxide (NO) composite insulator with different gate materials and subsequent anneals. Curve 1: Al gate; Curve 2: As *in situ* doped n$^+$-poly gate without anneals; Curve 3: As ion-implanted n$^+$-poly gate annealed (RTA) at 1000C for 6 sec. Current density = 1.93×10^{-5} A/cm^2.

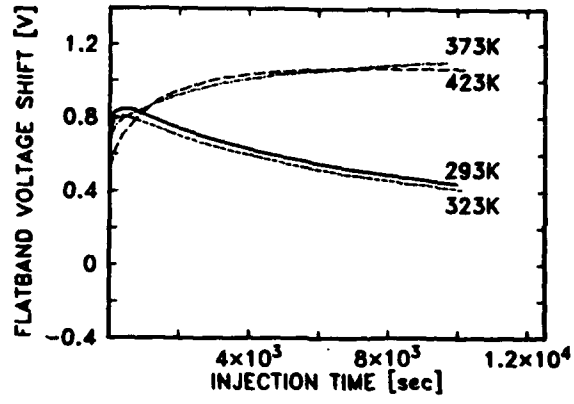


Fig. 3. V_{FB} shift as a function of injection time at different temperatures for the 4 nm nitride + 8 nm oxide (NO) composite insulator with As ion-implanted n$^+$-poly gate annealed (RTA) at 1000C for 6 sec. Current density = 1.93×10^{-5} A/cm^2.

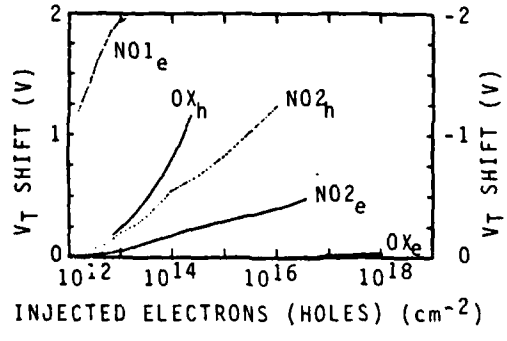


Fig. 4. V_T shifts due to electron and hole traps in FETs with various insulators measured by a substrate hot electron (hole) injection method: NO1 = 7 nm nitride + 8 nm oxide; NO2 = 4 nm nitride + 8 nm oxide; OX = 10 nm SiO_2. The polysilicon gate electrode was annealed at high temperature ($\geq$ 900C) Subscript e = electron injection; subscript h = hole injection.

ELECTRON AVALANCHE INJECTION IN THIN NITRIDED SiO_2 FILMS

Maurizio SEVERI, Maurizio IMPRONTA and Marco BIANCONI

CNR - Istituto LAMEL, Via Castagnoli 1, 40126 Bologna, Italy

Electron avalanche injection has been used to study charge trapping in thin (10-30 nm) ammonia-annealed silicon dioxide films as a function of process conditions. While electron traps (with a cross section of $\sim 10^{-16}$ cm^2) increase with nitridation temperature and time, the generation of interfacial positive charge ("turn-around effect") is greatly reduced or totally eliminated under severe nitridation conditions.

1. INTRODUCTION

Ammonia-annealed (nitrided) silicon dioxide films have been recently investigated as an alternative to thermal SiO_2 for thin gate insulators and for tunnel dielectrics /1/. The replacement of oxygen by nitrogen atoms brings about many modifications of the electrical properties of the films, which are closely related to the spatial distribution of the nitrogen. This, in turn, depends on the nitridation conditions /1,2/, postnitridation treatments and initial oxide thickness. It has been reported that the nitridation improves several characteristics of the insulator /3,4/. However, the electron trapping in nitrided films increases as compared with conventional oxides /5,6/. On the other hand, there is disagreement concerning the characteristics and the origin of these traps. Capture cross section of both $1x10^{-17}$ /5/ and $1x10^{-14}$ cm^2 /6/ have been reported from avalanche injection experiments. In /5/ the traps were ascribed to an increase of the OH content in the nitrided films. Moreover, a detailed study of the generation of interfacial positive charge during electron avalanche injection ("turn-around effect" /7/) in nitrided oxides is still lacking. In this work we present a trapping characterization of thin (10-30 nm) nitrided SiO_2 films upon electron avalanche injection as a function of process conditions.

2. EXPERIMENTAL

Al gate MIS capacitors were fabricated on p-type <100> silicon wafers of $(1-2)x10^{17}$ cm^{-3} doping. For 10 nm films, boron implanted substrates were used to obtain the optimum surface impurity concentration ($\sim 5x10^{17}$ cm^{-3}) for uniform injection /8/. The B^+ ion implantation conditions used were $6x10^{12}$ cm^{-2} at 20 keV + $1.5x10^{13}$ cm^{-2} at 70 keV. Oxidation was performed in a standard furnace at 900°C in dry O_2. SiO_2 layers of 10-30 nm were grown and in-situ annealed at the same temperature in N_2 for 10 min. The oxide thickness was measured by ellipsometry. Nitridation was carried out at temperatures between 800-1100°C in ultrapure ammonia gas at atmospheric pressure in a cold wall RF-heated reactor. Some samples received a postnitridation annealing in N_2 at 1000°C or in O_2 at 950°C for 15 min to obtain ONO films. Al-gate capacitors were formed by e-gun evaporation through shadow mask with 0.6 mm dots. After back-oxide stripping, Al was deposited for substrate contact. Finally, the samples received a post-metallization anneal in N_2 at 450°C for 20 min.

Electrons were injected from the substrate into the dielectric using the avalanche technique /9/. The surface was driven into deep depletion by applying a 30 kHz sawtooth voltage. A feedback circuit keeps the average injected current at a constant value (10^{-7} A).

The oxide charge buildup was monitored by periodically interrupting the injection and performing a fast C-V measurement. From the shift of the flatband voltage V_{FB} with time, the electrical trap parameters were determined by computer analysis of the data by assuming that electron trapping is a first-order process and that no detrapping takes place.

3. RESULTS AND DISCUSSION

The effect of the nitridation conditions on the charge trapping is shown in Figs.1 and 2 for 30 nm SiO_2 films. The flatband voltage shift ΔV_{FB} is reported as a function of the injection time for films that were annealed in ammonia at various temperatures for a fixed time (30 min) (Fig.1), or for different times at a fixed temperature (900°C) (Fig.2). As expected, a large increase in electron trapping is observed after nitridation. The density of the electron traps increases with nitridation temperature or time. The average nitrogen concentration also increases for more severe nitridation conditions /1,2/. Moreover, the dominant electron trap has a capture cross section of $\sim 10^{-16}$ cm^2, which is larger than that attributed to water-related centers in oxide (10^{-17}-10^{-18} cm^2). Therefore, this trap does not seem to be related to OH or H_2O defects, but it is probably due to nitrogen itself. In fact, it has been shown that group V impurities (N,P,As,Sb) can act as electron traps when they occupy oxygen sites /10/.

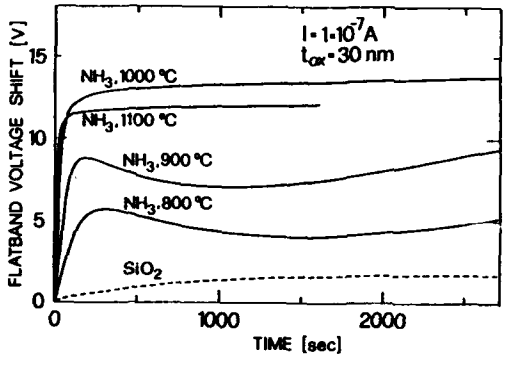


FIGURE 1

Effect of nitridation temperature on the charge trapping by electron avalanche injection at room temperature.

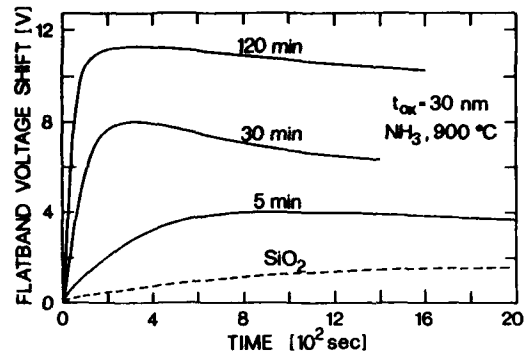


FIGURE 2

Effect of nitridation time on the charge trapping. $I = 1 \times 10^{-7}$ A.

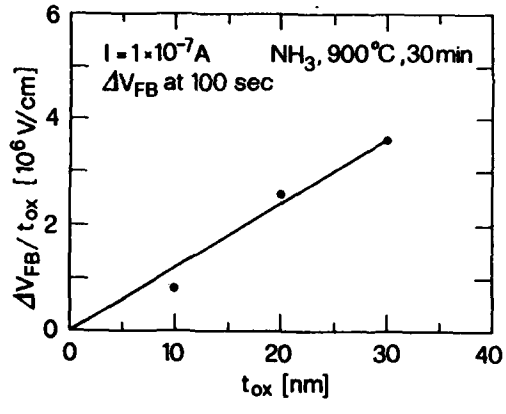


FIGURE 3

$\Delta V_{FB}/t_{ox}$ vs t_{ox} for films nitrided at 900°C.

Fig.3 shows ΔV_{FB} (taken after 100 sec) divided by the oxide thickness, t_{ox}, for samples with different t_{ox} nitrided at 900°C for 30 min. The trapping rate varies approximately as the square of the oxide thickness, which suggests that the traps are uniformly distributed throughout the nitrided oxide. A similar conclusion has been reached in /5/ on

the basis of photo I-V measurements.

The most interesting results of this work, however, are those related to the turn-around effect. As can be seen in Figs.1 and 2, the generation of positive charge at the insulator-silicon interface may be present also in nitrided oxides depending on the nitridation conditions. Only under severe nitridation conditions ($\geq$ 1000°C or $\geq$ 120 min) the positive charge generation is greatly reduced or eliminated. Notice that this happens in films with high density of electron traps. This is quite surprising, since the generation of the interfacial positive charge in conventional oxides has been usually correlated with electron trapping /7/. Apparently, there are some competing effects which control the generation of positive charge in nitrided oxides. At low nitridation temperatures or for short nitridation times, the generation of positive charge is more efficient than in standard oxides due to the increase of the electron traps and the lowering of the hole barrier at the metal-SiO_2 surface (which has been shown to increase the turn-around effect /11/). For more severe nitridation conditions, however, the nitrogen concentration near the interface increases /1,2/ at such a level that the dielectric structure becomes more resistant to interfacial damage. The mechanism by which this happens is not clear. We can only speculate, as suggested in /4/, that nitrogen, due to different bonding requirement, may impede the relaxation after breaking of strained Si-O bonds. The effect of nitrogen could also be due to changes in the film stress which can affect the interface state generation by a modification of the density of the strained Si-O-Si bonds near the interface /12/. In fact, it has been shown that it is possible to tailor the N concentration in deposited oxynitrided films in such a way to obtain film and substrate stress-free /13/. It is interesting to underline that a reduction of positive charge generation along with an increase of electron trapping has been recently reported for HCl oxides /14/ and it has been suggested that the presence of chlorine at the interface would reduce the oxygen dangling bond density due to the formation of chlorine-oxygen bonds.

We found that the generation of fast interface states during the avalanche injection, as revealed by the deformation of quasi-static C-V curves, is also greatly reduced in nitrided films which do not show the turn-around effect. A similar reduction in the generation of radiation-induced interface states has been reported for nitrided films that have been through severe nitridation conditions (1100°C, 6h) /15/. Nitridation also reduces the density of high field generated traps /4/. All these effects are probably related to the presence of a sufficient amount of nitrogen in the strained layer near the interface.

The dependence of the turn-around effect on the initial oxide thickness further supports the role of the nitrogen near the interface.

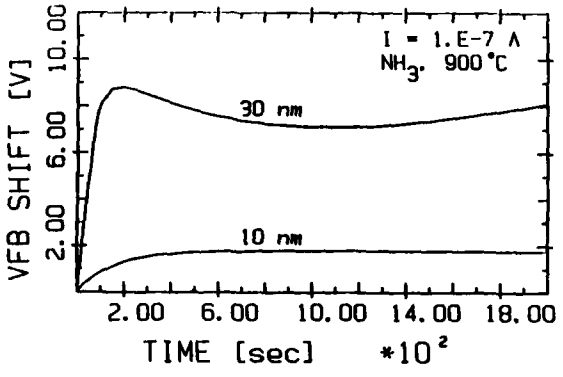


FIGURE 4

ΔV_{FB} as a function of electron avalanche injection time for 30 and 10 nm films nitrided at 900°C for 30 min.

Fig.4 shows the charge trapping behaviour of 30 and 10 nm films nitrided at 900°C for 30 min. The positive charge generation is practically eliminated in the 10 nm film. A qualitative connection with the fact that nitridation

of thicker films results in a lower concentration of nitrogen in the interfacial region /2/ may be made. Notice that in this case the suppression of the positive charge generation is also favoured by the concomitant reduction in the electron trapping.

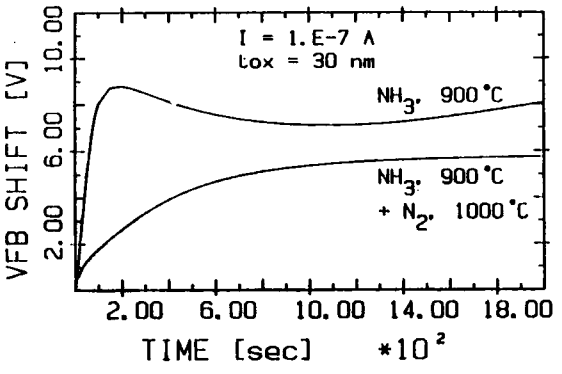


FIGURE 5

Effect of postnitridation annealing in N_2 for 120 min on the charge trapping in 30 nm films nitrided at 900°C for 30 min.

A similar effect has been observed for 30 nm films after a postnitridation annealing at 1000°C in N_2 for 120 min. The results are shown in Fig.5. The annealing both reduces the electron trap density by an order of magnitude and eliminates the turn-around effect. Annealing in O_2 is less effective in decreasing the trap density. Further investigation is necessary to fully clarify the role of the postnitridation treatments.

4. CONCLUSIONS

The electron trapping and the positive charge generation during avalanche injection through thin nitrided oxide films were studied as a function of the process conditions. The presence of a high density of electron traps, resulting from nitrogen at oxygen sites throughout the oxide, does not impede the generation of positive charge to be suppressed under severe nitridation conditions. The process dependence of the turn-around effect suggests that the nitrogen near the interface is the main responsible for this effect.

ACKNOWLEDGEMENTS

We wish to thank P.Negrini, G.Pizzochero, S.Guerri and P.Castelli for their assistance in sample preparation.

REFERENCES

/1/ Moslehi, M.M. and Saraswat, K.C., IEEE Trans.Electron Devices ED-32 (1985) 106

/2/ Habraken, F.H.P.M., Kuiper, A.E.T., Tamminga, Y. and Theeten, J.B., J.Appl. Phys. 53 (1982) 6996

/3/ Ito, T., Nakamura, T. and Ishikawa, H., IEEE Trans.Electron Devices ED-29 (1982) 498

/4/ Lai, S.K., Lee, T. and Dham, V.K., in IEDM Techn.Dig. (Washington, DC) (1983) 190

/5/ Lai, S.K., Dong, D.W. and Hartstein, A., J.Electrochem.Soc. 129 (1982) 2042

/6/ Chang, S.T., Johnson, N.M. and Lyon, S.A., Appl.Phys.Lett. 44 (1984) 316

/7/ Feigl, F.J., Young, D.R., Di Maria, D.J., Lai, S.K. and Calise, J., J.Appl.-Phys. 52 (1981) 5665

/8/ Dori, L., Arienzo, M., Nguyen, T.N., Fischetti, M.V. and Stein, K.J., J.Appl. Phys. 61 (1987) 1910

/9/ Young, D.R., Irene, E.A., DiMaria, D.J., De Keersmaecher, R.F. and Massoud, H.Z., J.Appl.Phys. 50 (1979) 6366

/10/ Pantelides, S.T., Thin Solid Films 89 (1982) 103

/11/ Fischetti, M.V., Weinberg, Z.A. and Calise, J.A., J.Appl.Phys. 57 (1985) 418

/12/ Zekeriya, V. and Ma, T.P., J.Appl.Phys. 56 (1984) 1017

/13/ Rand, M.J. and Roberts, J.F., J.Electrochem. Soc. 120 (1973) 446

/14/ Chen, A.J., Dadgar, S., Hsu, C.C.H., Pan, S.C.S. and Sah, C.T., J.Appl.Phys. 60 (1986) 1391

/15/ Terry, F., Aucoin, R., Naiman, M. and Senturia, S., IEEE Electron Dev.Lett. EDL-4 (1983) 191

Session B3.2

Compound Semiconductors Technology III

Chairman: G.F. Piacentini

Wednesday, September 16, 1987

DEPOSITION and PASSIVATION PROPERTIES of PLASMA CVD AlN FILMs on GaAs using METALORGANIC Al SOURCE

Fumio HASEGAWA, Tsuyoshi TAKAHASHI, Kiyokazu KUBO, Seinosuke OHNARI*, Yasuo NANNICHI and Toshiaki ARAI*

Institute of Materials Science, *) Institute of Applied Physics,
University of Tsukuba, Tsukuba Science City, 305 JAPAN.

Deposition conditions of plasma CVD AlN films were optimized, and the deposited films were examined as passivation films for the heat treatment of GaAs. It was found that when the AlN/GaAs was annealed at 850°C for 15 min. in H_2 atmosphere, the film was more oxidized than when it was annealed in Ar atmosphere. Correponding to the oxidation, the surface carrier concentration is more decreased for the AlN/GaAs sample annealed in H_2 atmosphere. Raman spectroscopy indicated that quite amount of stress was induced at the interface of AlN/GaAs when it was annealed in H_2, and TO phonon which is inhibited on (100) surface was observed. When the AlN/GaAs was annealed in Ar atmosphere, the Raman peak shift due to the stress and the TO phonon was not observed. These results suggests that the PCVD-AlN is a good passivation film of GaAs when it is annealed in Ar atmosphere.

1. INTRODUCTION

III-V compound semiconductors represented by GaAs have superior properties to Si such as a high electron mobility, direct band gap and possibility of heterojunctions. However, one of the biggest drawbacks of the compound semiconductors is that a good insulating or passivation film, such as SiO_2 for Si, has not been available so far. Therefore, the GaAs MISFET has not been successful. Passivation films for annealing after the ion implantation still have some problems such as Ga out diffusion into SiO_2 film or stresses between SiN_x film and GaAs[1,2].

On the other hand, there are some wide bandgap III-V materials which might be possible to be used as an insulator or as a passivation film for GaAs. AlN seems to be most suitable material because it has a bandgap of 6.2 eV and the same expansion coefficient as GaAs. Actually, it is reported that sputtered AlN could be used as a good passivation film of GaAs for the annealing after ion implantations[3]. However, plasma CVD generally gives less damage on III-V materials than the sputtering.

In the previous paper, we demonstrated for the first time that amorphous AlN film can be deposited by the plasma CVD[4]. In this paper, we would like to present the first report on the passivation properties of these PCVD-AlN films.

It was found that if the deposition conditions are optimized and the annealing is performed in Ar atmosphere, the PCVD AlN films can be used as a good passivation film of GaAs.

2. DEPOSITION SYSTEM and DEPOSITION CONDITIONS

2.1. Deposition system

A commercially available, parallel electrode plasma CVD system, ANELVA PED-301, was used by modifying the gas control unit. The schematic diagram of the deposition system is shown in Fig.1. The TMA --$Al(CH_3)_3$-- was supplied to the chamber through the upper electrode with H_2 carrier gas. The NH_3 was supplied from the side wall of the chamber in a separate gas line from the TMA. When they are mixed in the same gas control unit, they react to make something white and stuffs the gas line as described in the previous paper[4].

The temperature of the TMA bubbler and the flow rate were changed from 16'C to 25'C and from 10 to 50 cc/min., respectively, but typically the bubbler temperature of 16'C and flow rate of 40 cc/min. were adopted. The NH_3 flow rate was changed from 10 to 60 cc/min. Pressure of the chamber was kept at 1.0 - 2.0 Torr. The substrate temperature was typically 300'C, and the rf power was varied from 20 to 250W.

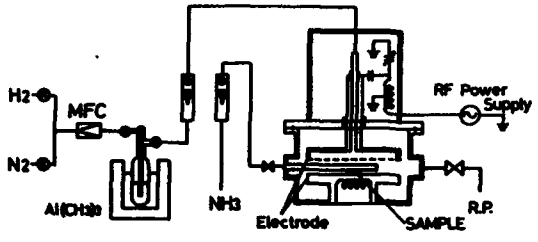


Fig.1, Plasma CVD system for the AlN deposition with TMA and NH_3.

2.2. Dependence on the TMA and NH_3 supply

Figure 2 shows the dependence of the deposition rate, etching rate and refractive index on the supply of TMA. The etching rate was measured by 30'C phosphoric acid (H_3PO_4), and the refractive index was measured by an ellipsometry. The deposition rate and the etching rate increase with increase of the TMA supply. The refractive index decreases from nearly 2.0 to 1.85 with increase of the TMA flow rate from 20 to 50 cc/min.. There was no dependence of the deposition rate on the NH_3 flow rate. These results indicate that the deposition rate is determined by the supply of Al and a better film is obtained for a higher NH_3/TMA ratio.

3. PROPERTIES as a PASSIVATION FILM for GaAs ANNEALING

3.1. Samples and experiments

Si doped conductive LEC GaAs ($n=6\times10^{16}$ cm^{-3}) was used as the substrate to see change of the surface carrier concentration and the stress at

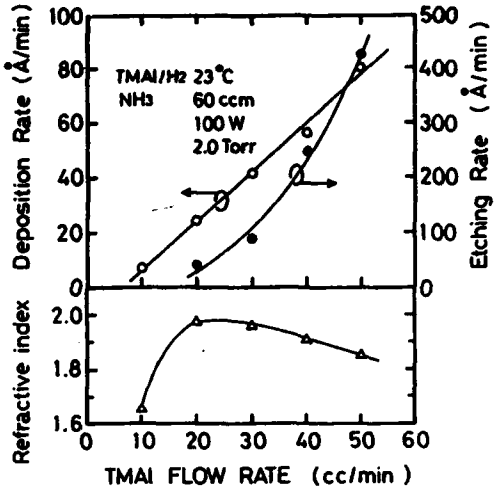


Fig.2, Dependence of the deposition rate, etching rate and refractive index on the TMA flow rate.

the interface. The AlN films were deposited on the GaAs substrates in nearly the optimum conditions (TMA/H_2;16'C/40cc, NH_3;60cc, rf power;20W, Ts;300'C), and were annealed at 850'C for 15 min. in H_2 or in Ar atmosphere. Composition profiles were measured by sputtering Auger Electron Spectroscopy (AES). Stress at the AlN/GaAs interface was estimated by Raman spectroscopy.

3.2. Change of the deposited AlN films by the heat treatment (annealing)

Figure 3 shows the composition profiles of an AlN/GaAs structure after the heat treatments in H_2 or in Ar atmosphere. The N content was about 2/3 of the Al content, but this film gave better passivation characteristics than the one whose Al/N ratio was 1.

When the sample was annealed in Ar, change of the composition profile was not detectable, but for the sample annealed in H_2, the oxygen content increased to about 10% and the N content decreased slightly. This result is contrary to a common sense, because H_2 is generally thought to prevent oxidation of the sample. One possibility is that oxygen in the H_2 gas converts to H_2O, and since H_2O molecule is smaller than O_2

molecule, the H_2O goes into the AlN film and oxidize the Al.

There is no indication of Ga or As out diffusion into the AlN film after the annealing, which is a serious problem of SiO_2 passivation films on GaAs. In that sense, the PCVD AlN film is better than at least SiO_2 as the passivation film of GaAs.

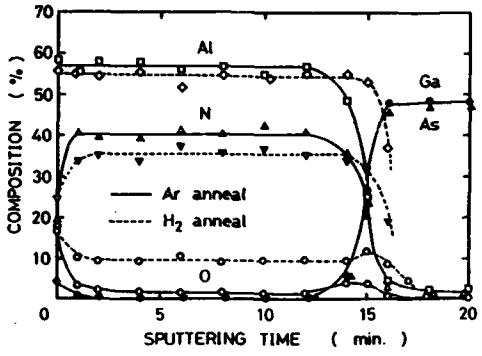


Fig.3, Composition profiles of the AlN/GaAs structure after the annealing in H_2 or in Ar atmosphere.

Figure 4 shows changes of the refractive index and the etching rate with 30'C H_3PO_4 as a function of the annealing temperature. The annealing time is 30 min.. The deposition conditions of the film are listed in the figure. The etching rate decreases very quickly with increase of the annealing temperature, and saturate at temperatures of more than 700'C. This fact should mean that the film becomes much denser by the heat treatment, but the refractive index does not seem to increase by the annealing. This is probably due to low accuracy of the measurement of the refractive indexes. There is no difference of the property change by different ambient gases during the annealing, even though the films were more oxidized in the H_2 atmosphere.

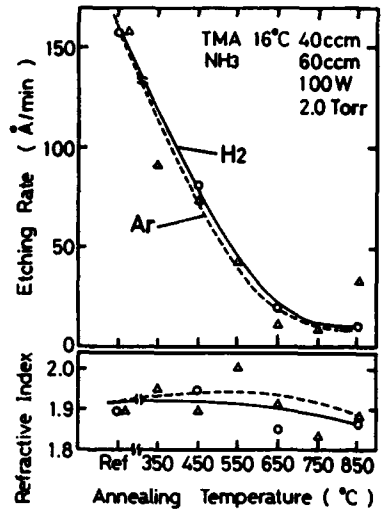


Fig.4, Dependence of the etching rate, refractive index on the annealing temperature.

3.3, Change of the carrier concentrations of the GaAs surface

Figure 5 shows the carrier concentration profiles of the GaAs surface before and after the heat treatment with the AlN passivation film. When the ambient gas is H_2, decrease of the surface carrier concentration is observed,

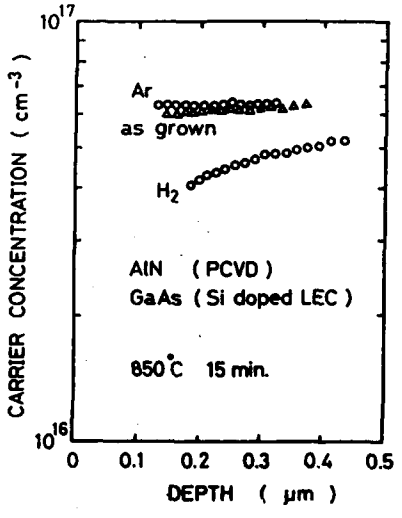


Fig.5, Change of the surface carrier density profile by the anneling in H_2 and in Ar.

but for the annealing in Ar, the carrier concentration profile is almost the same as that of the as grown one. This fact also indicates that Ar is a better atmosphere for the annealing of the AlN/GaAs structure.

3.4. Raman spectra of the annealed AlN/GaAs

It is reported that the stress at the passivation film-substrate interface can be estimated from the shift of Raman spectrum[5]. In order to see the stress at the AlN/GaAs interface, Raman spectra of the annealed samples in H_2 and in Ar atmosphere were measured. The 6471A line of Ar laser was used for the excitation on the (100) GaAs surface with the AlN film. The measurement was performed at room temperature.

The results are shown in Fig.6. The Raman shifts for the as-deposited sample and the one annealed in Ar ambient at 850°C for 15 min. were only the shift due to LO phonon, and it is 291.5/cm which is almost the same as the one for free GaAs surface (292/cm). The Raman spectrum of the sample annealed in H_2 atmosphere was different from those of the as-deposited one or the one annealed in Ar. A small signal of the shift due to TO phonon is observed at 266/cm. Since the Raman shift due to TO phonon is prohibited

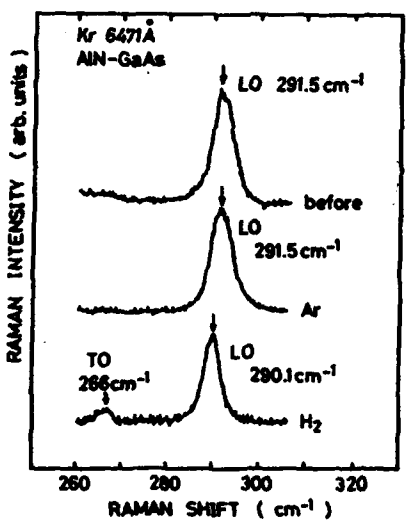


Fig.6, Raman spectra before and after the annealing in H_2 or in Ar.

on (100) GaAs surface, this fact means that the surface of this sample is slightly disordered by the heat treatment. Furthermore, the peak due to LO phonon is 290.1/cm, which is shifted from the as-deposited one by 1.4/cm. According to Cerdeira et al.'s results[6], this shift means that there is stress of about 10^{10} dyn at the interface. This large stress at the interface must be related with the facts that the AlN film is oxidized, and the surface carrier density is decreased by the annealing in H_2 atmosphere.

Acknowledgments

The authors are deeply grateful to Drs. S.Yoshida and S.Misawa of Electrotechnical Laboratories for valuable discussions.

This work was supported by the Scienctific Research Gran-in-Aid for Special Project research on "Alloy Semiconductor Electronics," from the Ministry of Education, Science and Culture of Japan.

References

[1] Konig,U. and Sasse,E. J. Electrochem. Soc., 130 (1983) 950.

[2] Hasegawa,F., Yamamoto,N., Sumi,M. and Nannichi,Y. Suppl. to Extd. Abst. of 15th Conf. on Solid State Devices and Materials, Tokyo, 1983, p48, C4-3LN.

[3] Okamura,S., Nishi,H., Inada,T. and Hashimoto,H., Appl. Phys. Letts. 40 (1982) 689.

[4] Hasegawa,F., Takahashi,T., Onodera,I. and Nannichi,Y. Extd. Abst. 18th (1986 Int.) Conf. on Solid State Devices and Materials, Tokyo, 1986, p663.

[5] Nakamura,T., Ushinokura,A. and Katohda,T., Appl. Phys. Letts. 38 (1981) 13.

[6] Cerdeira,F., Buchenauer,C.J., Pollak,F.H. and Cardona,M., Phys. Rev. B. 5 (1972) 580.

AN INVESTIGATION OF DEEP LEVELS IN GaAs FETs BY SELECTIVE DE-EXCITATION OF THE DEEP DONOR LEVEL EL2

J. MADDEN, M.R. BROZEL, A.R. PEAKER, G. ASHCROFT*

Department of Electrical Engineering and Electronics and the Centre for Electronic Materials, University of Manchester Institute of Science & Technology, P.O. Box 88, Manchester.M60 1QD. England.

A modified photo-FET technique is presented where the Idss photoresponse before and after EL2 quenching is measured. The technique has been applied both to ion-implanted FETs and to VPE grown devices with buffer thicknesses ranging from 0.7 μm to 3 μm grown on chromium-doped substrates. The removal of the EL2 contribution to the photoresponse leads to a modification of the charge states of the remaining deep levels whose presence can be detected by examining the difference between the quenched and unquenched Idss spectra. In the case of VPE FETs, a minimum buffer thickness to avoid chromium diffusion problems can be established.

1 INTRODUCTION

It has been known for some time that certain defects in GaAs FETs can be revealed by observing the photoresponse of the Saturated Drain Current at zero bias, I_{dss} [1]. It is generally accepted that such effects result from the extrinsic ionization of deep centres in the high resistivity region beneath the channel layer. In many cases the photo-emission of carriers from traps affects the channel width. Traps in the active region can also be studied by Deep Level Transient techniques such as DLTS and this allows centres to be studied and their concentrations estimated [2,3].

However, the photoresponse is often complicated by different centres which are excited simultaneously and the spectrum is difficult to interpret. In this paper we present a modification of the photo-I_{dss} technique where the dominant defect, EL2, is selectively and reversibly quenched. The photoresponse before and after quenching of the EL2 centres allows an analysis of the residual defect energy levels to be made.

2 EXPERIMENTAL

Amongst the deep level centres in GaAs FET structures, both epitaxial and ion-implanted is EL2. This native defect is believed to be a complex involving the arsenic antisite As_{Ga}.

2 EXPERIMENTAL

Amongst the deep level centres in GaAs FET structures, both epitaxial and ion-implanted is EL2. This native defect is believed to be a complex involving the arsenic antisite As_{Ga}.

EL2 centres exhibit the remarkable property of being rendered both electrically and optically inactive after irradiation by sub-bandgap light at temperatures below ~ 120K [4-6]. It follows that any photoresponse of GaAs FETs which involves EL2 can be modified if these centres can be "quenched" in this way. Additionally, because EL2 is a deep donor, modifications to the responses of other levels after quenching provides information about *their* charge states.

The experimental arrangement is indicated in Figure 1. The GaAs FET is held in a variable temperature (77-500K) cryostat. The device can be illuminated via quartz windows from a double grating monochromator and a quartz halogen bulb. The wavelength range is from

* Plessey (Caswell) Ltd., Towcester, Northamptonshire. England.

above the bandedge at 0.85 μm to 2 μm. I_{dss} is displayed as a function of wavelength on an X-Y recorder.

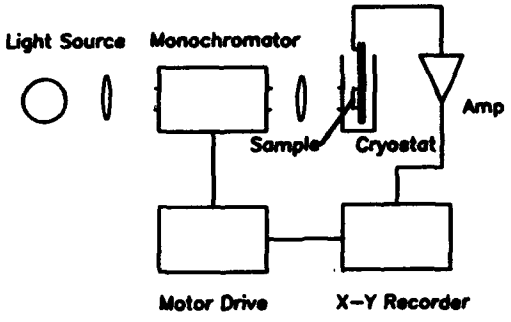


FIG 1 Experimental arrangement.Bleaching light source sits between monochromator and cryostat.

Initially, the device is cooled to approximately 80K. Changes in the spectral response of I_{dss} are found to occur at wavelengths shorter than about 1.6 μm, at longer wavelengths there is little or no response. This spectral response is totally reproducable and there is no evidence of changes being induced due to quenching of EL2. However, after the FET is illuminated with the strong light from a naked incandescent bulb, placed near the cryostat, the spectrum is modified.

The new spectral response is stable at temperatures below 100K, but recovers to the previous, non-quenched state once the temperature is raised above 120K. This behaviour is a well-accepted characteristic of the EL2 centre.

Figure 2 shows the 80K photoresponse before and after quenching. The FET is produced by direct Si^{29} ion-implantation into an undoped, semi-insulating LEC GaAs substrate. It is seen that not only can photo-excitation increase and decrease I_{dss} as a function of wavelength, but that in this case, quenching reduces the spectral response also as a function of wavelength.

Figure 3 shows a similar result obtained on a FET produced by VPE growth on a chromium-doped substrate. The technique used for these samples involved the growth of a

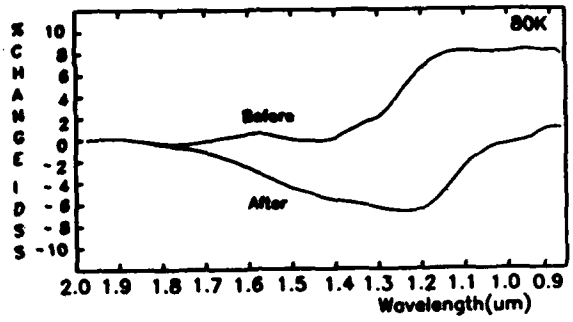


FIG 2 Variation of IDSS with wavelength for ion implanted FET,before and after bleaching.

thin, nominally undoped buffer layer ranging in thickness from 0.7 μm to 3 μm on a Cr doped substrate. A sulphur doped epi-layer channel was then grown, followed by a silicon doped contacting layer. In this case with a buffer thickness of 0.7 μm photo-quenching increases I_{dss} around 1.6 μm by nearly 60%

It was found that the photosensitivity of these FETs varied markedly with buffer thickness and this has been correlated to out-diffusion of Cr into the buffer layer [7-9].

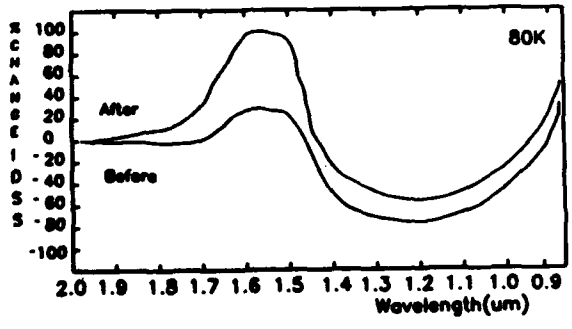


FIG 3 Variation of IDSS with wavelength for VPE FET before and after bleaching.

Previous work suggests that the buffer layer contains EL2 centres at concentrations of 10^{14} to $10^{15}cm^{-3}$ in addition to Cr atoms which have diffused from the substrate. Spectral changes after quenching are greater for VPE grown FETs with buffer layer thicknesses of less than 1 μm than the ion-implanted devices.

3 DISCUSSION

The photoresponse of GaAs FETs must be due to modifications of the channel conductivity

beneath the Schottky gate. Whether or not these are due to direct excitation within the depletion region, the pinched-off channel or elsewhere, extrinsic photoresponse must be due to carriers excited from deep levels. We can assume, like several other authors [1,10,11], that this photo-excitation takes place either in the buffer layer or in the region adjacent to the channel/buffer (for epitaxial FETs), or near the channel/substrate interface for ion-implanted FETs. In other words, we probe the deep level defects in regions essentially depleted of carriers and photo-excitation is, therefore, important.

Consider the band diagram in Figure 4, where it is assumed that deep acceptors in addition to EL2 are present; shallow levels are ignored. The allowed transitions which produce photo-excited carriers are indicated. Clearly, the removal of electrically active EL2 defects by quenching (Figure 4(b)) modifies the spectrum appreciably; the three ionization paths being replaced by a single transition.

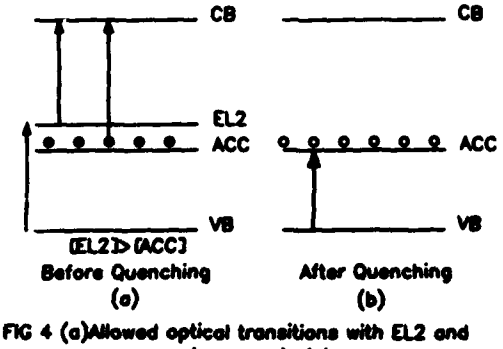


FIG 4 (a)Allowed optical transitions with EL2 and another deep level (acceptor), (b) allowed transitions with EL2 removed by bleaching.

In addition, the relative position of the deep acceptor level with respect to the EL2 level can also be deduced because if this was higher in the band gap than EL2, quenching would not affect its charge state.

The difference between the spectra before and after quenching reveals information on centres have been modified and on the EL2 centres which have been removed.

Figures 5 and 6 show the difference spectra for 0.7 μm buffer VPE FET and the ion-implanted FET respectively.

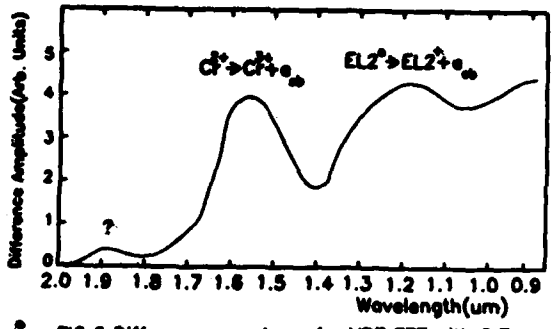


FIG 5 Difference spectrum for VPE FET with 0.7um buffer.

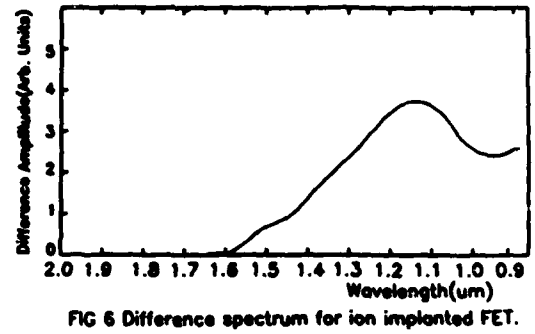


FIG 6 Difference spectrum for ion implanted FET. The spectrum shows the optical ionisation cross section for EL2,suggesting that no other centres were excited within the wavelength range.

The ion implanted difference spectrum shows that the photo-I_{dss} response was due to EL2 alone [12,13]. The VPE FET difference spectra are complicated, but they reveal the EL2 peak as before. The change in charge state of the Cr atoms caused by the removal of the EL2 is also revealed.

We have employed DLTS to confirm that chromium atoms are present in the channel layer and that our interpretation of the photo-I_{dss} spectra is reasonable. However, simple DLTS does not indicate different charge states whereas the method reported is sensitive to these parameters. The data have been used to estimate the best buffer thickness for VPE FETs. Figure 7 shows how the photo-sensitivity of the FETs varies with buffer

thickness. As the buffer thickness decreases to less than 2.5 μm the changes in I_{dss} (both positive and negative) become significant.

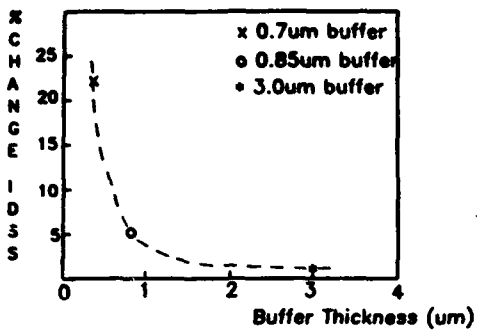


Fig 7 IDSS sensitivity to buffer thickness on Cr doped substrates.

4 CONCLUSIONS

It has been shown that the photoquenching of EL2 from FET structures can cause significant modifications to the photoresponse of these devices so revealing previously hidden processes. Positions of deep centres with respect to EL2 can be established and, together with DLTS, can be used to characterise them. The technique is non-destructive and can be used on standard devices.

REFERENCES

[1] Tegude. F.J.; Heime. K: Electron. Lett.16, 22 (1980)

[2] Adlestein. M.G.: Electron.Lett. 12, 297 (1976)

[3] Hawkins, I.D., Peaker, A.R., Appl.Phys. Lett. 48 (s), 20 Jan 1986 p 227-229

[4] Taniguchi.M.; Ikoma. T: J.Appl.Phys. 54, 6448 (1978)

[5] Lagowski.J.; Lin.D.G.; Aoyama. T.; Gatos. H.C: Appl.Phys.Lett. 44, 336 (1984)

[6] Yahatta. A.; Nakajima.M: J.Appl.Phys. 43 L313 (1984)

[7] Asbeck.P.; Tandom.J.; Sin.D.; Fairman.R; Welch.B: IEEE GaAs IC Symposium.Research Abstracts No 14 (1979)

[8] Evans. C.A.; Deline. V.R: IEEE GaAs IC Symposium.Research Abstracts No 14(1979)

[9] Forbes. L.; Chang.C.D: IEEE GaAs IC Symposium.Research Abstracts No 14(1979)

[10] Kocot.C.; Stolt.C.: IEEE Trans. on Elec. Dev. Vol. ED29 No. 7 (July 1982)

[11] Sriram.S.; Das.M.B: Solid State Electron Vol.28, No. 10, pp 979-989 (1985)

[12] Martin. G.M.: Appl.Phys.Lett. 39. (November 1981).

[13] Chang.C.D.; Forbes.L: Semi-Insulating III-V Materials Nottingham, p.329 (1980)

A DLTS INVESTIGATION OF VPE GaAs MESFETs

C. Ghezzi*, E. Gombia, R. Mosca and M. Pillan**

Istituto MASPEC del Consiglio Nazionale delle Ricerche, Via Chiavari 18/A
43100 Parma, Italy
*Dipartimento di Fisica dell'Universita' di Parma, Via M. D'Azeglio 85
43100 Parma, Italy
**TELETTRA S.p.A., Via Trento 30
20059 Vimercate (MI), Italy

Capacitance DLTS measurements have been performed in VPE GaAs MESFETs prepared on Bridgman Cr-doped and LEC undoped semiinsulating substrates. A band of electron traps was detected near the metal (gate)-semiconductor interface. Near pinch-off conditions, a positive capacitance signal was found to dominate the DLTS spectra in samples prepared on Cr-doped substrates. The feature of this positive capacitance transient have been analyzed and discussed.

1. INTRODUCTION

In short gate GaAs MESFETs, the DLTS (deep level transient spectroscopy) technique is commonly used in the conductive mode [1,2] since the gate capacitance is too small for an accurate detection of the capacitance transients. Conductance DLTS signals, whose sign is consistent with hole (minority carrier) emission processes, have been frequently observed to dominate the DLTS spectra [1-4]. However, at present, a common accepted idea is that the "positive" signature of conductance DLTS signals in short gate GaAs MESFETs is related to electron emission from surface states in the ungated surface source-gate and gate-drain access regions [5-7]. Only a few signals, such as the one corresponding to the ubiquitous EL2 electron trap, seem then to be doubtless related to true bulk traps in high-quality devices. In the light of these results, and taking also advantage of the fact that conductance and capacitance DLTS give similar results in large area devices [8], we have reconsidered a capacitance DLTS investigation in VPE GaAs MESFETs using enlarged contact patterns.

2. EXPERIMENTAL METHODS

The GaAs MESFETs were prepared by Telettra SpA. Both Bridgmann Cr-doped and LEC undoped semi-insulating GaAs substrates were used. The active channel was obtained by S-doping during the final growth stages of the variable thickness VPE layers (1-6 μm) grown by the $AsCl_3$ method at T=700 C. A surface layer 0.2 μm thick was then removed by etching before the deposition of the TiPdAu gate contact. The maximum electron concentration n_{max} in the active channel was $1.4x10^{17}$ cm^{-3}, but a few samples were prepared with significantly lower concentrations ($5\text{-}6x10^{16}$ cm^{-3}). The residual concentration $|N_D\text{-}N_A|$ of uncompensated shallow impurities in the undoped (buffer) VPE region was in the range of 10^{14} cm^{-3}.

The DLTS investigation was performed using both a lock-in system and a double boxcar averager for processing the transient capacitance signal. In the latter case, filling pulses as width as 100 ms could be easily employed. To have significant gate capacitances we used an enlarged contact pattern (fat FET) with a 100 μm gate lenght and 10 μm interelectrode spacings. During measurements, the source and drain contacts were connected together. By means of a standard C-V technique the electron concentration in the active channel was profiled as a function of both position and gate reverse voltage V_g. A suitable choice of V_g makes then possible to take DLTS spectra which are related in a controlled manner to different depleted regions.

3. RESULTS

The highly doped active channel was investigated by keeping V_g at suitable low values. For samples with n_{max}= $1.4x10^{17}$ cm^{-3} this could be achieved using V_g=-0.6 V (see plot 1 in fig.1b). The corresponding

DLTS spectrum (plot 1, fig. 1a), taken at a voltage pulse amplitude V_p=0.6 V, shows a broad band of electron traps with concentrations $N_t=2n_{max}(\Delta C/C)$ in the range 3-6x10^{14} cm^{-3}. The DLTS signal has a maximum at a temperature not far from the one (T=265 K) corresponding to the electron trap EL3 (E_c-E_t=0.575 eV). This band was found in all the samples and the spectrum did not change using different width tp of the filling pulse: more specifically, positive capacitance signals were not observed even for t=40 ms. It must be stressed here that, under the specified conditions,the depletion region W is limited to 0.11 μm and the electron traps responsible of the observed band are even closer to the gate contact. A more deep region (W=0.28 μm) in the active channel, although far from the pinch-off conditions, could easily be investigated in samples with a lower doping level (n_{max}=5-6 x 10^{16} cm^{-3}) using

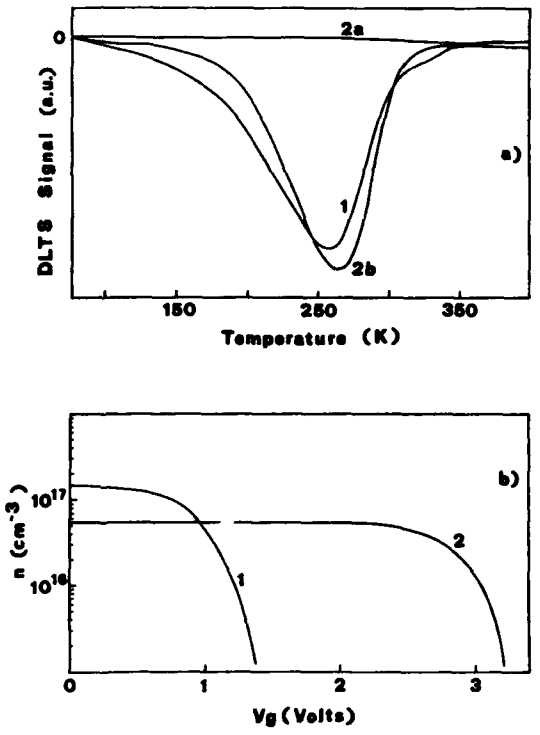


Figure 1

DLTS spectra (1a) within the conducting channels of MESFETs having different doping profiles (1b). Rate window=21.7 sec^{-1} (lock-in system). See text for further details.

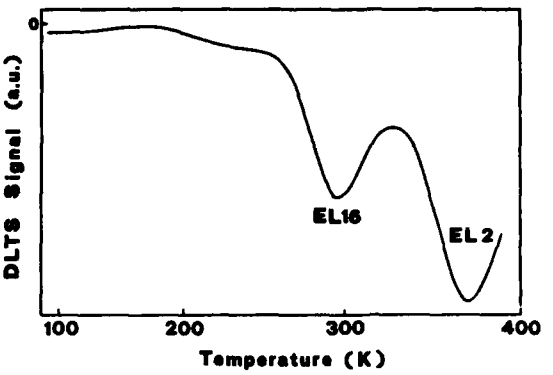


Figure 2

Typical DLTS spectrum of a GaAs MESFET prepared on a LEC undoped substrate. V_g=-1.8 V is near pinch-off conditions. V_p=0.6 V, rate window=49.5 sec^{-1} (boxcar system).

V_g=-2.2 V (plot 2, fig.1b). The occurence of a flat spectrum taken at V_p=1 V (plot 2a, fig.1a) means that the trap concentration within this region can be estimated to be smaller than 1x10^{13} cm^{-3}. As a check, if V_g is reduced to -0.2 V in the same sample, in order to investigate a shallower region (W=0.16 μm), a band of electron trap can be again observed (plot 2b, fig.1a). It should be noted that, owing to the different doping levels in the two conductive channels, the trap concentration corresponding to curve 2b is roughly 2.5 times smaller that the one of case 2a. We may conclude that the observed traps, being absent in the deep channel region, are not related to the VPE growth process. Moreover, they cannot be related either to the interruption of the VPE growth, since a 0.2 μm surface layer was removed. The fact that they could be related to defects introduced close to the contact by the gate metallization process deserves further investigations.

When even deeper regions were investigated and the active channel was nearly depleted, quite different results were obtained for MESFETs prepared with different substrates. For LEC undoped substrates both the electron trap EL16 (E_c-E_t=0.37 eV) and EL2 (E_c-E_t=0.78 eV) were detected (fig.2), while in the case of Cr doped substrates the DLTS spectra were dominated by the presence of a positive capacitance signal whose amplitude was strongly dependent on both the frequency

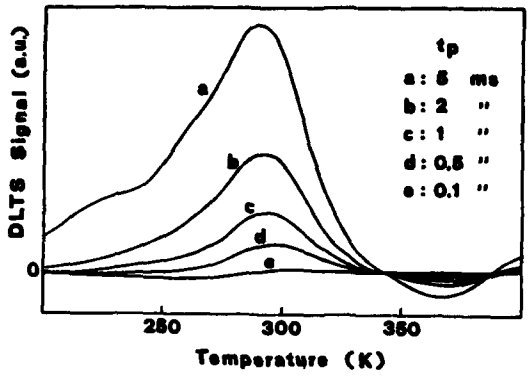


Figure 3

Typical DLTS spectra of a GaAs MESFET prepared on a Cr-doped substrate. V_g=-1.8 V, V_p=0.6 V, rate window=21.7 sec^{-1} (lock-in system). Different spectra refer to different width tp of the filling pulse.

and the filling pulse width t_p (fig.3). The signal amplitude saturates at t_p=50 ms. The origin of this signal can hardly be related to surface effects within the interelectrode spacings, similar to those reported in [5-7]. In fact it was completely absent in MESFET samples grown on LEC undoped substrates. Moreover it could be detected only when Vg was sufficiently large, not far from the pinch-off value (fig.4). The positive $\Delta C/C$ signal, if due to the variations of the trapped surface charge, for the simplified case of a homogeneus sample, is roughly given by

$$\frac{\Delta C}{C} \sim \frac{4D}{L_g+4D} \frac{\Delta W}{W}$$

with L_g=gate length, D=interelectrode spacing, $W=W(V_g)$=depletion width under the reverse gate voltage V_g and ΔW=variation of W due to the voltage pulse V_p. Since W is an increasing function of V_g and since ΔW decreases by increasing V_g, the amplitude of the positive signal should continuously decrease by increasing V_g. This is contrary to the observed behaviour.

Finally, with reference to the possibility that a high series resistance R_s may reverse the sign of the appearent Capacitance transient, as suggested in [9], we checked the R_s value in our samples by comparing low and high frequency values of the gate-drain conductance. Under operative conditions we derive $Q=Rs\,\omega C \sim 0.1$, which is far from the value (Q=1) for reversing the signal.

As a comment, the possibility of bulk hole traps being responsible for the observed positive DLTS signal might be reconsidered. It was impossible, however, to get reliable activation energies since two emission processes , at least, contribute to the positive DLTS signal, as argued from fig.4. As suggested by Zylberstejn et al. [2], the population of a hole trap level can be changed by pulsing the gate since the hole quasi-Fermi level F_p can intersect the level at the barrier between the active layer-substrate interface. We believe that this model could be still valid, in principle, even if the barrier is distributed on the whole buffer thickness as a consequence of a strong depletion effect. The hole density can then be small enough to be responsible for low capture rates. The role of Cr doped substrates should be considered by accounting that Cr diffusion can both introduce trap levels and change the shape of the barrier within the buffer layer.

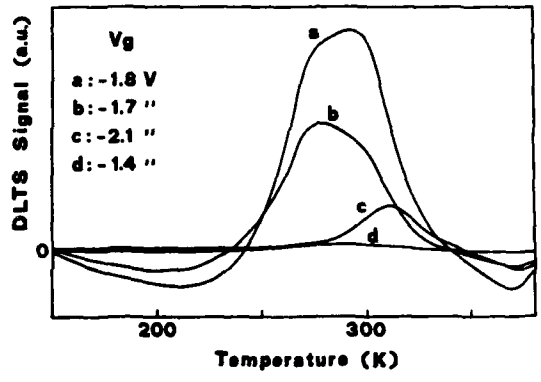


Figure 4

DLTS spectra of a GaAs MESFET prepared on a Cr doped substrate. The different spectra refer to different V_g values, as indicated. The filling pulse reaches the 0 V level. t_p=2 ms, rate window=21.7 sec^{-1} (lock-in system).

4. CONCLUSIONS

A capacitance DLTS investigation has been performed in GaAs MESFETs. Two main results can be summarized: (i) a band of electron traps has been detected near the interface between the Schottky gate

contact and the active channel. This band is not detected deep in the active layer, so it cannot be related to the VPE growth process (ii) positive capacitance DLTS signals whose features are hardly consistent with the hypothesis of surface states have been detected near pinch-off conditions.

REFERENCES

[1] Alderstein, M.G., Electron. Lett. 12 (1976) 297

[2] Zylberstejn, A., Bert, G. and Nuzillat, G., Inst. Phys. Conf. Ser. 45 (1978) 315

[3] Itoh, T. and Yamai, H., IEEE Trans. Electron Dev. ED-27 (1980) 1037

[4] Milonneau, A., Meignant, D., Baudet, P., Berth, M. and Hollan, L., Inst. Phys. Conf. Ser. 56 (1981) 445

[5] Wallis, R.H., Faucher, A., Pons, D. and Jay, P.R., Inst. Phys. Conf. Ser. 74 (1985) 287

[6] Blight, S.R., Wallis, R.H. and Thomas, H., IEEE Trans. Electron Dev. ED-33 (1986) 1447

[7] Harrang, J.P., Tardella, A., Rosso, M., Alnot, P. and Peray, J.F., J. Appl. Phys. 61 (1987) 1931

[8] Hawkins, I.D. and Peaker, A.R., Appl. Phys. Lett. 48 (1986) 227

[9] Broniatowsky, A., Blosse, A., Srivastava, P.C. and Bourgoin, J.C., J. Appl. Phys. 54 (1983) 2907

Session D3.2

Laser II

Chairman: J. Turner

Wednesday, September 16, 1987

ANALYSIS AND DESIGN OF PHASE-LOCKED DIODE LASER ARRAYS

F. LOZES-DUPUY, A. BENSOUSSAN, S. BONNEFONT, G. VASSILIEFF, H. MARTINOT

Laboratoire d'Automatique et d'Analyse des Systèmes du CNRS
7, Avenue du Colonel Roche - 31077 TOULOUSE CEDEX - FRANCE

A general eigenmode analysis is applied to model typical index-guided phase-locked laser arrays. It is shown that in phase operation can be achieved if only a very narrow range of structural parameters is controlled. CSP arrays are found to exhibit efficient supermode discrimination.

1. INTRODUCTION

Semiconductor laser arrays are receiving widespread interest due to their capabilities for high power applications. Such devices could be used in optical communications between satellites, short-haul local area networks, laser printers, optical recording systems.

One promising approach is to phase lock closely spaced narrow stripe lasers by evanescent coupling. The control of the array modes, however, is the main problem to be solved. Most laser diode arrays oscillate in the out-of phase lateral mode, yielding a double-lobe far-field pattern.

In this paper, a coupled-mode analysis is applied to describe the operation of phase-locked injection laser array. The model is applied to typical index laser arrays in order to define structures operating with all the elements in phase and having a resulting narrow single-lobe far-field pattern.

2. COUPLED-MODE ANALYSIS

Coupled-mode analysis [1,2] is used to analyse the lateral modes of the arrays. The N supermodes supported by an array consisting of N weakly coupled single-mode lasers are given by solving the well-known set of the simultaneous equations :

$$\left.\begin{array}{l}(\beta^L-\beta_p)A_p^L + K_{p,p-1}\,A_{p-1}^L + K_{p,p+1}\,A_{p+1}^L = 0 \\ p = 1, \ldots, N \text{ and } L = 1, \ldots, N\end{array}\right\} \quad (1)$$

The eigenvalue β^L and the eigenvector $(A_1^L, \ldots, A_N^L)$ are respectively the propagation constant and the amplitudes of the individual elements of the array operating in the L th supermode.

The propagation constants β_p of the isolated lasers and the coupling coefficients K are to be first determined in order to solve the eigenvalue problem for the modes of the array. A simple and relatively good estimate can be made by treating each individual laser as a three-layer slab waveguide along the y-direction parallel to the junction planes. The stripe and interstripe regions form in the x-direction multilayer dielectric waveguides with effective refractive indices [3] corresponding to a fundamental TE-like mode.

The model is capable of analysing both uniform and chirped arrays. It takes into account the active and passive nature of the layers by using complex refractive indices, including carrier-induced index change relations [4]. Thus the model should be useful to simulate several kinds of index-guided laser arrays, various geometries or nonuniform current distributions.

3. ANALYSIS OF TYPICAL LASER ARRAYS

Following the model, the main parameters of an array are the number of elements, the stripe width, the stripe spacing, the lateral optical confinement of an individual laser.

On the other hand, single-lobe far field pattern and efficient discrimination between the supermodes require very restrictive conditions. In the case of uniform arrays, the modal gain of the L th supermode is given by

$$G^L = G - 4 K_i \cos \frac{L\pi}{N + 1} \qquad (2)$$

where G is the modal gain of the isolated individual laser, N the number of elements, K_i the imaginary part of the coupling coefficient. Single-lobe far field pattern associated with the fundamental supermode is achieved if $K_i<0$, and good discrimination between the fundamental and the highest-order supermode is provided by large values of $|K_i|$.

The model has been applied to typical index guided laser arrays in order to compare their capability for stable single beam operation.

Buried heterostructures are found to present strong index guiding, eliminating coupling between the lasers and providing broad far field pattern.

Arrays consisting of ridge waveguide lasers, channeled substrate planar (CSP) structures, V-channeled substrate inner stripe (VSIS) lasers are also considered. They are shown in cross-section in Fig. 1. In each array case, the layer thickness t is a key parameter which determines the lateral optical confinement for a given set of structural parameters.

Calculations were made with y=0.40, x=0.05 (λ = 0.84 µm), d = 0.1 µm. The real values of the refractive indices are n = 3.641 in the active layer (without pumping), n = 3.343 in the cladding layers, n = 3.630 in the GaAs layers. Internal loss are taken as α = 10 cm^{-1} in the cladding layers, α = 8000 cm^{-1} in the GaAs lossy layer. Gain in the active layer is adjusted until the array modal gain reaches 50 cm^{-1}.

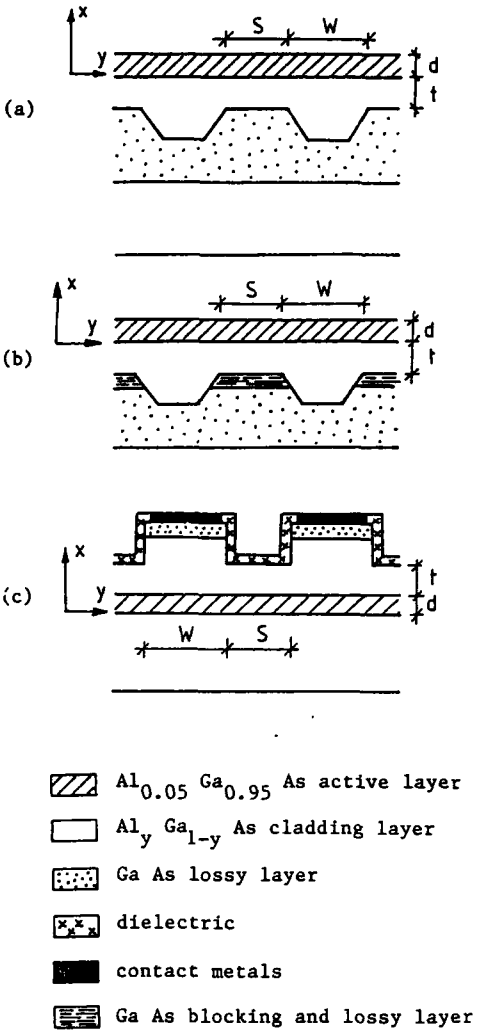


FIGURE 1

Schematic diagram of phase-locked arrays composed of (a) CSP structures, (b) VSIS lasers, (c) ridge waveguide lasers.

10-emitters arrays with stripe width W=2 µm and separation S = 1 µm are analyzed. The thickness t is determined so that such arrays favour the fundamental supermode over the anti-phase supermode (K_i < 0). The corresponding lateral optical confinement of the individual laser is shown in Fig. 2. It can be seen that the index-guiding of the VSIS array is quite poor leading to lateral mode instabilities and

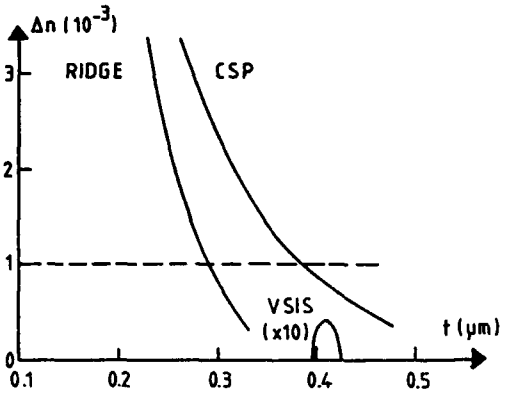


FIGURE 2

Lateral optical confinement of individual laser versus layer thickness t for the case of CSP, VSIS and ridge arrays operating in the fundamental supermode.

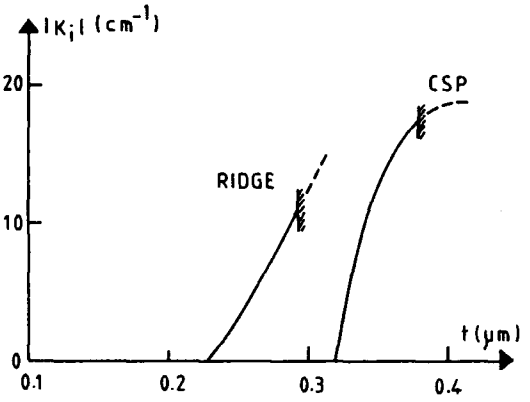


FIGURE 3

Amplitude of the imaginary part of the coupling coefficient versus layer thickness t for the case of CSP and ridge arrays operating in the fundamental supermode.

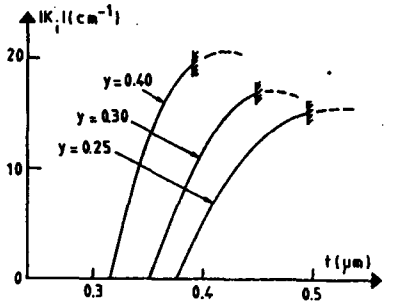


FIGURE 4

Amplitude of the imaginary part of the coupling coefficient versus layer thickness t for CSP arrays with different compositions of the cladding layers.

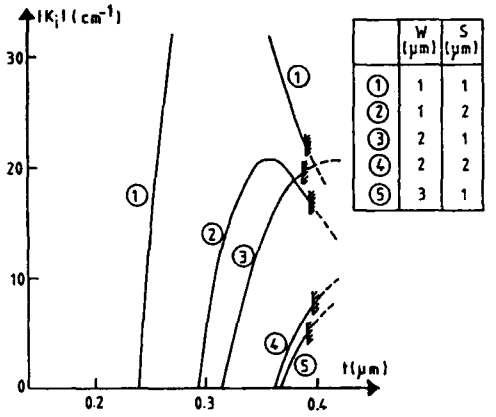


FIGURE 5

Amplitude of the imaginary part of the coupling coefficient versus layer thickness t for CSP arrays with different stripe and interstripe widths.

similar properties to that of a broad-area device. Ridge and CSP arrays provide stronger optical confinement. The amplitude of the imaginary part of the coupling coefficient, shown in Fig. 3, indicates that both structures may achieve modal gain discrimination by choosing the layer thickness in a very narrow range of values, with slightly better potential in the case of the CSP array. Note that the maximum value of $|K_i|$ -and therefore the maximum modal gain discrimination- is limited by a constraint of a minimum lateral optical confinement of the individual laser which has been taken equal to 10^{-3}. CSP arrays have an additional advantage over the ridge structure since stable waveguiding mechanisms are maintained when the pumping current is increased.

Other design constraints can be imposed by fabrication limits such as composition and doping material uniformity, or variation in the layer thicknesses. Fig. 4 gives the available range of the layer thickness t of the CSP ar-

rays when the composition of the cladding layer is modified. Fig. 5 shows that close spacings and narrow stripes are required to achieve fundamental mode oscillation. These calculations are consistent with published experimental results.

4. CONCLUSION

A coupled-mode analysis applied to typical arrays, indicates that fundamental supermode operation can be provided by a careful choice of structural parameters. Good modal discrimination between the fundamental and highest-order supermode is found for CSP arrays. Moreover mode stability is achieved by these structures because they little suffer from the importance of the effects of carrier-induced dielectric constant changes. However the main problem to be solved remains the homogeneity of the wafer and the limits of the fabrication in order to insure uniform growth with well controlled parameters.

ACKNOWLEDGEMENTS

The authors acknowledge contract support from the European Space Agency and the Centre National d'Etudes Spatiales.

REFERENCES

[1] J.K. BUTLER, D.E. ACKLEY, D. BOTEZ
Appl. Phys. Lett. 44 (1984) 293.

[2] E. KAPON, J. KATZ, A. YARIV
Opt. Lett. 10 (1984) 125.

[3] W. STREIFER, E. KAPON
Appl. Opt. 18 (1979) 3724.

[4] J. MANNING, R. OLSHANSKY, C.B. SU
IEEE J. Quantum Electron QE-19 (1983) 1525.

REDUCTION OF LOSSES IN GAIN-GUIDED LASER ARRAYS COMPARED WITH SINGLE STRIPE LASERS

J. Gerner and W. Schairer
TELEFUNKEN electronic
D-7100 Heilbronn

INTRODUCTION

Coherently coupled laser arrays have found rapidly increasing interest during recent years. One reason for this is their capability of very high output power.

Because of their dramatically reduced losses compared with single stripe lasers, laser arrays resemble in some respect broad area stripe lasers and therefore can be simpler treated theoretically than narrow single stripe lasers. Thus the quality of the epitaxial process can be more directly judged from the analysis of arrays than of single stripe lasers. On the other hand, there is usually a wealth of empirical data connecting the quality of the epitaxy with the properties of single stripe lasers.

We deemed it therefore useful to compare the properties of arrays with those of single stripe lasers, which have been prepared on the same wafer and therefore have the same epitaxial structure, saw the same wafer processes and have the same mirror quality.

EXPERIMENTAL

The epitaxial layers have been prepared by liquid phase epitaxy. Three layers form a double hetero structure (DH) and a p-type GaAs top layer is used as a contact layer as shown in detail in Fig. 1.

Oxide stripe lasers are prepared using a very shallow Zn-diffusion to facilitate the formation of low resistivity stripe contacts. Contact metallizations are Ti/Pt/Au on the p-side and Ni/Au:Ge/Ni/Au on the n-side. On each wafer one, five and ten stripe lasers are fabricated and the pitch of the stripes is varied between 5, 6, 9 and 12 µm. The processed wafers are cleaved into bars, usually 250 µm long.

RESULTS

Experimental results obtained from these GaAlAs multi and single stripe DH-lasers are summarized in the Table. The data collected refer to unpassivated laser chips. An assymmetric Al_2O_3 passivation would increase the differential efficiency η by about a factor 1.6. It is important to notice that the scatter of the measured data is quite low on each single wafer, but also between wafers. E.g. typical variations in threshold current I_{th} and differential efficiency η are only $\pm$ 5 %. This is also true for the emission wavelength shift $\Delta\lambda$ between single and multi stripe lasers, the scatter in the emission wavelength λ, however, is much lower.

The threshold current density I_{th} for the single stripe laser has been corrected by the subtraction of that part of the current which diffuses in the active layer in lateral direction

and is lost for the pumping of the active stripe region.

In the calculation of the nominal threshold current density J_{nom} the efficiency of the stimulated emission has been assumed as 0.7. Similar values have been obtained experimentally, but are also reported in the literature /1/. The calculated differential efficiency η_{calc}, the threshold current density J_{calc} and the wavelength shift $\Delta\lambda_{calc}$ are, however, not very sensitive to the choice of η_{stim}.

The nominal threshold current density can now be used within Stern's model /2/ to calculate the maximum gain g_{max} and the wavelength shift. The agreement between experimental and theoretical results is surprisingly good.

A test of the general consistency of the data can be performed by using g_{max}, which has been obtained within Stern's model, to calculate the internal loss α_s. The loss in turn may be used to calculate threshold current density J_{calc} and differential efficiency, η_{calc} according to the formulas

$$J = \frac{d}{\eta_{stim}} 4.5\cdot10^3 \frac{A}{cm^2\mu m} + \frac{d}{\eta_{stim}\Gamma} 20 \frac{A}{cm\mu m}\left\{\alpha_s + \frac{1}{L} \ln\left(\frac{1}{R}\right)\right\} \qquad (1)$$

$$\eta = \left\{\eta_{stim}/(1 + \alpha_s L/\ln\left(\frac{1}{R}\right))\right\} 1.24/2\lambda\ \mu mV \qquad (2)$$

These calculated results are also given in the Table. The threshold current density agrees closely with the experimental results whereas the differential efficiency approaches the experimental values only for the ten stripe devices. Nevertheless we judge the overall consistency as generally satisfactory.

CONCLUSIONS

The following conclusions are derived from this generally consistent picture:

(1) The properties of the laser arrays are determined by the quality of the epitaxial process and not deteriorated by an unsufficient accuracy of the realization of the array structure.

(2) The low threshold current density obtained for arrays proves that the threshold current density of narrow stripe lasers is as low as can be expected for this type of structure.

(3) The quantitative understanding of the wavelength shift results in a design rule connecting the Al content of growth melts with the emission wavelength.

Die diesem Bericht zugrundeliegenden Arbeiten wurden mit Mitteln des Bundesministers für Forschung und Technologie, FKZ 13N5399/8, teilweise gefördert. Die Verantowrtung für den Inhalt liegt jedoch allein bei den Autoren.

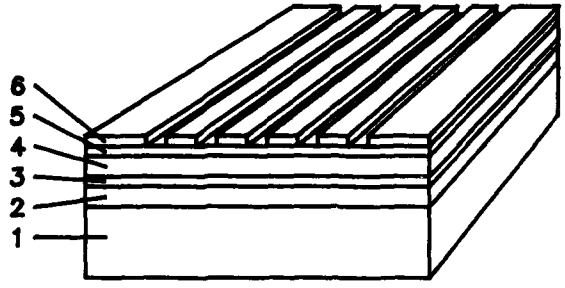
REFERENCES

/1/ E. Pinkas, B.I. Miller, I. Hayashi, and P.W. Foy, J. Appl. Phys. 43, 2827 (1972)

/2/ F. Stern, J. Appl. Phys. 47, 5382 (1976)

Fig. 1

Layer structure of single and multiple stripe DH lasers



Layer	Composition	Doping	Thickness
1 n–substrate	GaAs	Si: 10^{18} cm^{-3}	80µm
2 n–cladding	$Al_{.4}Ga_{.6}As$	Te: $5*10^{17}$ cm^{-3}	2.5µm
3 active layer	$Al_xGa_{1-x}As$	none	0.07...0.10µm
4 p–cladding	$Al_{.4}Ga_{.6}As$	Mg: $1*10^{17}$ cm^{-3}	1.5µm
5 p–cap	GaAs	Mg: $1*10^{17}$ cm^{-3}	0.5µm
6 oxide isolation			0.12µm

TABLE

Experimental and calculated results on single and multiple stripe DH lasers. Resonator length is in all cases 250 µm and the lasers are not passivated.

No. of stripes	1	5		10	
Stripe pitch/µm	-	5	6	5	6
$F/10^{-5}$ cm^2	1.25	6.25	7.5	12.5	15
I_{th}/mA	62	100	110	180	200
η/W/A	0.33	0.49	0.49	0.49	0.49
J/KA/cm^2	3.45	1.6	1.5	1.4	1.3
d/µm	0.1	0.1	0.1	0.1	0.1
J_{nom}/KA/cm^{2}µm	24.2	11.2	10.5	9.8	9.3
λ/nm	807	813	813	814.5	814.5
$\Delta\lambda$/nm	-	6.0	6.0	7.5	7.5
g_{max}/cm^{-1}	985	335	300	265	240
$\Delta\lambda_{calc}$/nm	-	6.3	6.9	7.5	8.0
α_s/cm^{-1}	220	45	35.4	26.0	19.2
J_{calc}/KA/cm^2	3.45	1.60	1.50	1.40	1.33
η_{calc}/W/A	0.092	0.27	0.30	0.34	0.38

LOW THRESHOLD BH LASERS EMITTING AT 1.5 μm MADE FROM GAS SOURCE MBE HETEROSTRUCTURES

B. FERNIER, L. GOLDSTEIN, D. BONNEVIE, D. SIGOGNE, J. BENOIT *
LABORATOIRES DE MARCOUSSIS - CR.CGE, Route de Nozay, 91460 MARCOUSSIS, FRANCE

C. CARRIERE, T. LAVOLEE
ALCATEL/CIT, Route de Nozay, 91460 MARCOUSSIS, FRANCE

Buried heterostructures (BH) lasers are made from low threshold current density (1.7 kA/cm^2 at 1.5μm) GaInAsP/InP double heterostucture (DH) grown by Gas Source Molecular Beam Epitaxy (GSMBE) and using a standard LPE regrowth fabrication process. Low threshold current (25 mA), high quantum differential efficiency (43 %) BH lasers emitting at 1.5μm have been obtained. Due to the very good homogeneity in thickness and composition of the epitaxial layers over a 2" diameter GSMBE wafer, the dispersion of the useful laser parameters is very low (± 6 mA for threshold current, ± 5 % for differential efficiency, ± 3 nm for wavelength). These lasers exhibit a stable behaviour on preliminary aging test.

1. INTRODUCTION

High quality heterostructures with a good homogeneity in thickness and composition over large area are needed either to improve the production yield of standard lasers or to allow the fabrication of more advanced devices such as DFB lasers for coherent systems, optical amplifiers or integrated optoelectronic circuit. Regarding this problem GSMBE appears as a very suitable technique as it will be shown in the case of GaInAsP/InP heterostructures for standard 1.5μm lasers.

2. GSMBE GROWTH OF GaInAsP/InP

The use of gas sources (AsH_3, PH_3) for V elements has already been recognized [1],[2] to be a very efficient way for the MBE growth of uniform and high quality $Ga_xIn_{1-x}As_yP_{1-y}$ ($y \simeq 2.2\ x$) layers lattice matched to InP. For the heterostructures discussed below the general conditions for GSMBE growth can be described as follows : arsine and phosphine are introduced into the growth chamber through a low pressure PBN cracking cell heated at 1000°C ; Ga and In fluxes are provided by solid effusion cells. The growth temperature is 540°C. In these conditions, layers with a good homogeneity in thickness ($\Delta t/t < \pm 5\%$) and composition ($\Delta\lambda g/\lambda g < 10^{-3}$) and a low background concentration ($n \simeq 2.10^{15}$cm^{-3}) can be grown on 2" diameter substrates in the whole composition range of the GaInAsP alloy. The low growth rate (3.5 Å/s) allows an accurate control of the layer thickness.

3. DESIGN AND CHARACTERIZATION OF LASER HETEROSTRUCTURES

Several types of heterostructures with (SCH) and without (DH) separate optical and carrier confinements suitable for low threshold laser operation at 1.3μm and 1.5μm have been designed following a previously published model [3], grown in the conditions described above, and characterized by photoluminescence, capacitance profiling and measurements on broad area laser.

3.1. Design of laser heterostructures

Conventionnal DH with a 0.15-0.2μm thick active layer and SCH with a 0.12μm thick active layer sandwiched between two 0.08μm thick guiding layers with higher phosphorus content ($\lambda g \simeq 1.3\mu m$) have been designed and grown for 1.5μm laser emission. The heterostructures consist of a four (DH) or a six (SCH) layer structure grown on a (100) S doped N type substrate ($n \simeq 6.10^{18}$ cm^{-3}) with a 2μm thick Sn doped ($n \simeq 10^{18}$ cm^{-3}) InP buffer layer, the guiding and active layers, a 2μm thick, Be doped ($p \simeq 3.10^{17}$ cm^{-3}) InP confinement layer and a 0.3μm thick, Be doped ($p \simeq 4.10^{18}$ cm^{-3}), quaternary ($\lambda g = 1.3\mu m$) contact layer. The thicknesses and

* This work was supported in part by EEC. ESPRIT project 263 A : "InP based optoelectronics"

doping level of the layers are controlled by SEM measurement and capacitance profiling.

3.2. DH characterization by photoluminescence

Room temperature photoluminescence (P.L) under direct excitation (λ = 1.06μm) of the active layer is used to check the quality and the homogeneity of the heterostructure. Together with a narrow P.L spectrum ($\Delta E \simeq 55$ meV), a low dispersion ($\Delta\lambda < 2$ nm) of the peak wavelength is observed on two inches wafers, indicating a good compositional homogeneity. As shown on Fig. 1, the low fluctuation of the P.L intensity over 2" wafers, which is typically lower than 10 %, allows to predict a low dispersion in the laser characteristics. A dispersion of the mean peak P.L wavelength, from wafer to wafer, lower than 4 nm is indicative of the reproducible control of the active layer composition.

3.3 Measurements on broad area lasers

Threshold current density (J_{th}) lower than 2 kA/cm^2 are reproducibly obtained on 400μm long broad area lasers issued from 20 wafers with DH or SCH emitting at 1.3μm or 1.5μm. As shown in table 1, values of J_{th} of 1.7 kA/cm^2 with a minimum value of 1.5 kA/cm^2 are currently obtained for 1.5μm DH with a 0.15μm thick active layer. As predicted theoretically [3] lower J_{th}, mean value 1.6 kA/cm^2 with a minimum value of 1.3 kA/cm^2, are obtained for 1.5μm SCH with a 0.12μm thick active layer. In addition to a low dispersion of their threshold current, the broad area lasers characteristics (Fig. 2) are reproducible and linear above threshold.

The external quantum differential efficiency depends on the internal optical losses ; the typical measured value η = 34 ± 4% for 400μm long 1.5μm lasers from the different GSMBE DH is significantly higher than for LPE DH ($\eta \simeq$ 25%). Such a difference is attributed to lower optical scattering losses due to a better interface flatness.

The good reproductibility (4nm) from run to run and the low dispersion (for each wafer) of the lasing wavelength (3nm) can also be observed in table 1.

Table 1 : Threshold current density and lasing wavelength for broad area lasers from several wafers with different active layer thicknesses (d).

N°	Struct.	d (μm)	J_{th} (kA/cm^2)	λ (nm)
2204*	DH	0.17	1.8 ± 0.15	1530 ± 3
3400	DH	0.2	1.93 ± 0.2	1571 ± 2
3456	DH	0.2	1.95 ± 0.13	1570 ± 2
3468	DH	0.19	1.92 ± 0.1	1563 ± 1.2
3478	DH	0.2	1.81 ± 0.05	1244 ± 1
3479	SCH	0.12	1.50 ± 0.12	1558 ± 1.4
3480	DH	0.15	1.65 ± 0.17	1575 ± 2
3491	SCH	0.12	1.70 ± 0.12	1561 ± 3.5
3492	SCH	0.12	1.58 ± 0.14	1568 ± 2.2
3512**	DH	0.15	1.62 ± 0.12	1546 ± 1.1

* All the structures are grown in the same MBE reactor except sample 2204.
** This sample correspond to a different choice for the active layer composition.

4. BH LASER (λ = 1.5μm)

4.1. BH fabrication

Following a standard fabrication process [4,5,6] which involves the LPE regrowth (T = 600°C) of burying layers, several GSMBE DH have been used in order to fabricate 1.5μm BH lasers.

The good thickness uniformity of GSMBE BH facilitates (Fig. 3) the control of the position of the neck of the etched mesa regarding the active layer which determines the active stripe width (< 2μm) and the most important laser characteristics [5].

4.2. BH characteristics

For each processed wafer, the L.I characteristics under pulsed conditions are measured on 50 unselected 250μm long chips lasers. For all the wafers 90% of the lasers have threshold current below 60 mA. Fig. 4 shows an histogram of the threshold current for a typical wafer, with a 35 mA mean value and a ± 6.5 mA standard deviation.

BH lasers have been mounted on standard copper heat sinks, bonded with In junction down, for DC measurements. A typical L.I DC characteristic linear up to four times the threshold current (24.5 mA) is shown on Fig. 5. Such a laser shows a stable transverse fundamental mode emission up to ≃ 15 mW without any facet coating. Fig. 6 illustrates the excellent temperature behaviour of these lasers which can emit up to 3 mW at 100°C. The measured external quantum efficiency 43± 5 % (L = 250μm) is in good agreement with measurements on broad area lasers. The homogeneity of the active layer and the high quality of the interfaces of the GSMBE DH has been further confirmed from spectral and near field measurements. The spectral properties of the BH lasers are characterized by a low dispersion of the lasing wavelength (eg λ = 1528 ± 3.2nm) and a reproducible regular

spectrum with few longitudinal modes and a red shift of the peak wavelength when the power is increased up to 10 mW (Fig. 7a). The subthreshold spectrum reveals a uniform gain profile (Fig. 7b) even for lasers with a longer cavity (L = 800μm) which are usefull for special devices such as high gain TW amplifiers [7] or narrow linewidth DFB laser [8]. The emitted power is characterized up to 10 mW by very stable and clean far field patterns (Fig. 8) suitable for coupling with a fiber or an external cavity.

4.3 Preliminary aging tests

A hard screening test performed at 80°C during 72 h at constant DC current of 150 mA has been performed on 10 BH lasers. If one except a few number of lasers showing a rapid increase of the threshold current in the first 24 h, most (70%) of the lasers do not show any appreciable increase (within 1%) of the 20°C threshold current. A long term aging test has also been performed on a first group of lasers at 7 °C at constant emitted power (5 mW). After 800 h the relative increase of the 20°C threshold current is lower than 10% for 70% of the samples. These first results, to be completed, indicate that no special failure mechanism relevant to GSMBE growth can be detected.

5. CONCLUSION

GSMBE DH wafer have been used to fabricate BH laser using a conventionnal LPE regrowth process. These lasers are characterized by a low threshold current (25 mA), a high efficiency (43%) and well controlled spectrum and radiation pattern. The low dispersion of the laser characteristics specially the lasing wavelength ($\Delta\lambda$ = 3 nm) and their stable behaviour during preliminary aging test indicate that GSMBE can be used for the production of standard 1.5μm lasers. Furthermore the above quoted performances make GSMBE very suitable for the fabrication of advanced components and IOEC's.

ACKNOWLEDGEMENTS

The authors acknowledge M. BOULOU, M. BOURDON, L. LE GOUEZIGOU, O. LE GOUEZIGOU, Y. LOUIS, C. PADIOLEAU, F. POINGT, and S. WEBER for helpful cooperation.

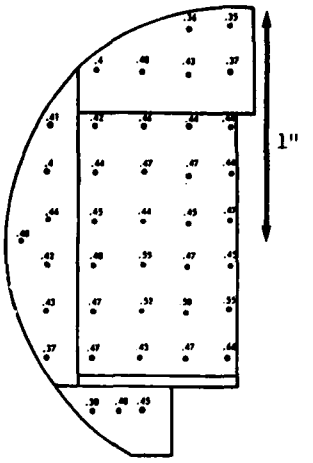


FIGURE 1

Photoluminescence intensity mapping of the active layer of a DH (a.u.).

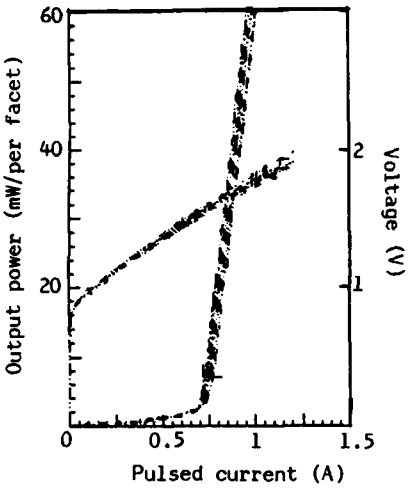


FIGURE 2

Light current characteristic of 9 unselected broad area lasers issued of a 1.3μm DH wafer.



FIGURE 3

SEM observation of an InGaAsP laser issued from a MBE double heterostructure with LPE regrown burying layers.

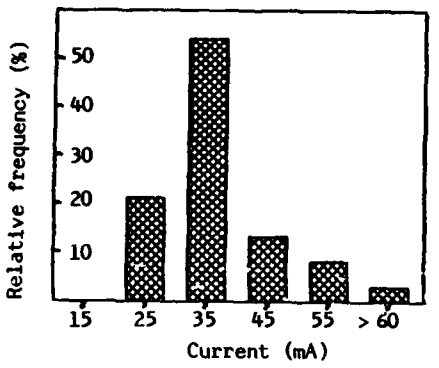


FIGURE 4

Histogram of threshold current of 50 BH lasers from the same wafer.

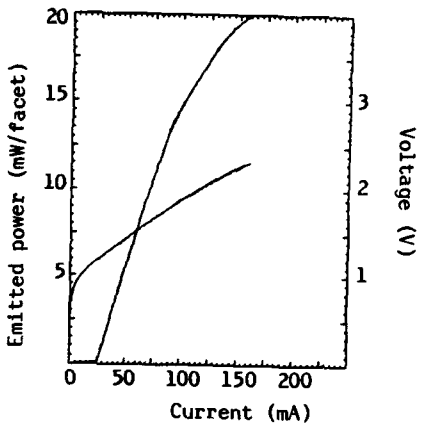


FIGURE 5

Light and voltage current characteristics under DC injection at 20°C of a 1.5μm BH laser.

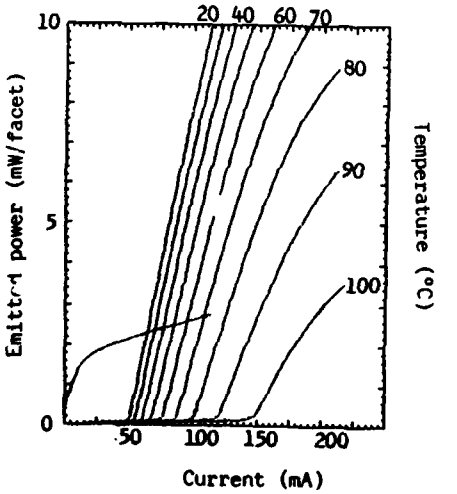


FIGURE 6

DC characteristics of a 1.5μm BH laser operating up to 100°C bonded p side down.

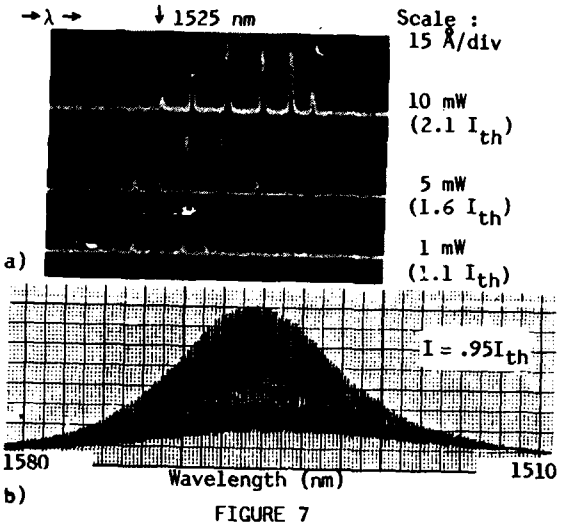


FIGURE 7

a) Longitudinal mode spectra of a 250μm GSMBE/LPE BH laser as a function of power.
b) Spectral gain profile of a 800μm GSMBE/LPE BH laser biased under threshold.

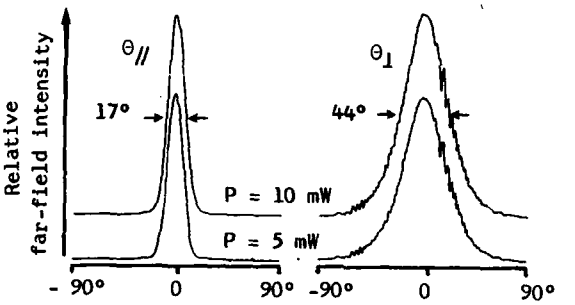


FIGURE 8

Far field patterns in the directions parallel and perpendicular to the junction plane for 5 and 10 mW emitted power.

REFERENCES

[1] Panish M.B., et al, J. Vac. Sci. Tech., B-3 (2), pp. 657-665 (1985).
[2] Huet D., et al, J. Electron. Mat., 15, 37 (1986).
[3] Fernier B, et al, IEE Proc. J. Optoelectron., 134, (1), pp. 27-34 (1987).
[4] Benoit J., et al, Proc. 9th European Conference on Optical Communication (ECOC'83), Geneva, Switzerland, pp. 35-38 (1983).
[5] Fernier B., et al, Proc. 3rd European Fiber Optic Communications and Local Area Networks Conference (EFOC/LAN/85), Montreux, Switzerland, pp. 81-87 (1985).
[6] Oguey C., et al, Proc. OFC'85, San Diego, USA (1985)
[7] Brosson P., et al, Electron. Lett., 23 (6), pp. 254-256 (1987).
[8] Liou K.Y. et al, Appl. Phys. Lett., 50 (2), pp. 489-491 (1987).

IP.12. - INVITED PAPER

INTEGRATED OPTICS : $LiNbO_3$ OR SEMICONDUCTORS ?

M.PAPUCHON

THOMSON-CSF/LCR - Domaine de Corbeville - BP 10 - 91401 ORSAY (France)

1. INTRODUCTION

During the last several years, integrated optics research has rapidly evolved in two main directions. In the first, $LiNbO_3$ based waveguides by Ti in diffusion were used to demonstrate various components from low loss high speed modulators or matrix switch arrays to specific chips for fiber sensors like the fiber gyroscope. As a matter of fact, several of these devices become to be commercially available from different companies. In the second, waveguides are fabricated in semiconductor materials and rapid progresses in the growth techniques led to high quality devices in both GaAs and InP systems leading to realistic previsions for complete monolithic integration.

2. THE STATE OF THE ART

$LiNbO_3$ integrated optics is now investigated since more than ten years. As a matter of fact, the now standard in diffusion techniques have been used for the first time in this class of material in 1974 (1, 2) and one year after, the first directional coupler in $LiNbO_3$ was reported (3). Since that time numerous improvements were achieved resulting in a mature technology with some devices now on the market in small quantities.

The interest in $LiNbO_3$ waveguides lies on the relatively easy related technology, their electrooptic activity and the excellent properties of the fabricated devices. For example, losses of the order of .3 dB/cm at 1.3 μm are obtained by Titanium in diffusion and command voltages around few volts are common for the electrooptically active circuits. One of the first area of application is in the data transmission systems. In that field, broadband modulators are required and to achieve moderate driving voltages together with the largest bandwidths, travelling wave type devices are preferred to lumped circuits. As the crystal is a good dielectric, the electrodes are of the coplanar type leading to an interaction between the optical modes and an inhomogeneous electric field. To build such modulators, different electrode configurations can be used (symmetric or assymetric coplanar...) and their exact position with respect to the waveguides depends on the electric field component to be maximized for the interaction. For a directional coupler switch/modulator, the assymetric coplanar configuration is in general preferred. To achieve a characteristic impedance of 50 ohms, the ratio between the width of the narrow section and the spacing between the electrodes should be around .6 in $LiNbO_3$. As the electrode spacing is of the order of 10 microns, it is thus necessary to use thick metallizations to minimize the losses of the electrical transmission line. Thicknesses around 3 microns are easily obtained by gold electroplating and

are commonly used in high speed devices. For a MACH ZEHNDER configuration, symetric transmission lines are in general used and more than 15 Ghz bandwidth has been reported (4). In these type of modulators, the bandwidth is limited by the speed difference between the optical and modulating electrical waves. This phenomena imposes a trade off between the bandwidth and the driving voltages. To compensate this velocity mismatch a periodic interaction can be used by designing suitable electrodes (5) resulting in a bandpass circuit the center frequency of which can be choosen by varying the electrode periodicity. It has been shown recently (6, 7) that aperiodic configurations could be very attractive as they permit very large bandwidths with long interactions. Using these principles, a bandwidth greater than 20 Ghz (#40 Ghz theoretical) has been demonstrated for a MACH ZEHNDER modulator (half wave voltage : 11 V, interaction length : 10 mm) at 1.3 microns (8).

Because of the progresses made in $LiNbO_3$ technology it is now possible to envisage the fabrication of complex chips integrating several optical functions. Althought the integration level will always be limited by the finite length of the electrooptical devices to be able to achieve reasonnable driving voltages very interesting results have been obtained so far. A typical example is the recently demonstrated (8) optical switching matrix where 64 directional couplers have been integrated to form an 8 x 8 strictly non blocking switching system. Fabricated on a 60 mm long $LiNbO_3$ substrate it was composed of switches the interaction length of which was 2 mm. Consequently, the driving voltages at the 1,3 μm working wavelength were of the order of 30 volts. This device represents the highest degree of integration realized to day using $LiNbO_3$ technology.

When dealing with integrated optic chips, an important problem is the fiber to fiber insertion loss. In particular, the coupling losses should be minimized. As a matter of fact, driving voltages and coupling losses are not independent as low drive voltages can be obtained with highly confined optical modes while low coupling losses between single mode fibers and integrated optic waveguides can be achieved if the modes of the two optical guides are well adapted. In general, a compromize must be accepted between these two characteristics. One way to overcome the problem is to be able to obtain different mode sizes in the "active" and coupling regions. Recently such a principle has been applied to MgO/Ti diffused $LiNbO_3$ waveguides (9). MgO/Ti diffusion was used in the in and out coupling regions of a $LiNbO_3$ optical waveguide in order to adapt the single mode fiber mode to the optical waveguide one while preserving the confinement parameters in the electrooptic interaction region. By this method, 1,1 dB total insertion was achieved for a straight phase modulator compared to 2,6 dB for the same device without the MgO while preserving the same command voltage for the device. These results are very important as they indicate that future progresses will lead certainly to integrated optic devices presenting total insertion losses below 1 dB.

$LiNbO_3$ technology not only led to high performance devices but also to state of the art telecommunication systems demonstrations. As a typical example a long distance system experiment (1 Gbt/s over 150 km) (10) employed a Ti : $LiNbO_3$ waveguide modulator. The field of application of integrated optic technology is not limited to optical communication. As a matter of fact, interferometric optical fiber sensors are ideal systems to use the particularities of integrated Optic techniques. It is particularly true for well known fiber optic gyroscope where planar techniques not only lead to a size reduction of the optical system but also to the possibility of using "active" components to help in the sensor signal processing. High performance compact devices have been built using Ti:$LiNbO_3$ waveguide chips (11) gathering beam splitter, polarizer, phase modulators and means to minimize back reflexions. To day, $LiNbO_3$ based integrated optics is considered

to be a mature technology. However, it lacks the possibilities of achieving monolithic integration of lasers, modulators, detectors and eventually electronic circuitry. This is why a large amount of work is also devoted to the use of semiconductor type substrates. In that case the waveguides are more often fabricated using epitaxy techniques such as Molecular Beam Epitaxy (MBE), Liquid Phase Epitaxy (LPE) or Metal Organic Vapor Phase Epitaxy (MOVPE). The different refractive indices needed to achieve waveguiding are in general obtained using heterostructures of materials with different composition (GaAs/GaAlAs, InP/GaInAs or GaInAsP...).

For many years the main drawback for their use in integrated optic devices were the relatively high losses associated with the resulting waveguides.

Recently, decisive progresses were achieved both in the GaAs and InP systems.

Concerning classical semiconductor films, losses between 0,2 and 0,5 dB/cm were demonstrated in MOCVD grown GaAs single heterostructures channel guides (12), . 7dB/cm in MOCVD grown GaAs double heterostructures waveguides (13) and very recently losses as low as 0,4 dB/cm were reported in InP/InGaAsP MOVPE waveguides at 1,5 μm. These results are particularly important as for the first time, the losses in semiconductor based optical waveguides are of the same order of magnitude than with those obtained in $LiNbO_3$.

When using semiconductor materials, very particular structures can be utilized. For example one can employed ultra thin films forming Multiple Quantum Wells (MQW). In that way a specific refractive index can be synthetised by adjusting the well/barrier thickness ratio. Using these techniques low loss GaAs/GaAlAs (MQW) waveguides were fabricated using MOVPE (14) with values of 0,15 dB/cm at 1,52 μm.

The same method can be employed with InP based structures leading to InGaAs/InP (well/barrier) composed waveguides with losses around 0,8 dB/cm at 1,52 μm (15). It is important to note that, in addition to these new low loss structures, even if the classical electrooptic effect can be used to fabricate modulators, switches... semiconductors offer much more than the simple equivalent of $LiNbO_3$ circuits. As a matter of fact, the injection or depletion of carriers in suitably biaised diodes could lead to very attractive effects on both the real and imaginary part of the refractive indices. These effects could lead to electrically active components with very short length (< 1 mm) as was demonstrated (16) in InGaAsP/InP carrier injection switches where less than 100 mA were necessary to achieve switching with a measured cross talk of - 20,5 dB.

In addition to the "passive" interesting properties of MQW to synthetize artificial refractive indices their unique properties due to the carrier confinement in the wells can be applied to realize extremely efficient electrically controlled devices.

For example, phase modulators with phase shifts as high as 12° deg/V.mm has been reported in InGaAs/InP waveguides (17) and highly efficient electroabsorption MQW modulators can be obtained (18) because of the carrier confinement which leads to strong excitonic line with a very sharp peak near the band edge.

As a final step, semiconductor materials would permit complete monolithic optoelectronic integration. As a matter of fact, the studies have already begun mainly in the field of integrating electronic components with Lasers or Photodetectors. (For a review, see (19)). The main difficulty is obviously the technological compatibility of the different circuits to be integrated. As a typical example, the electronic circuits (like FETs) are in general fabricated on semiinsulating (SI) substrates while optoelectronic components use currently conductive ones. It is then necessary either to learn how to fabricate opto devices on SI substrates (leading to the so call horizontal structure) or to be able to grow high quality insulating layers on conductive material (leading to the so call vertical structure). One of the main goal in that field is to achieve compact optical transmitters

or receivers for optical telecommunications. The experimental work has been mainly concentrated on GaAs system with demonstrations (among others) such as a transmitter integrating two lasers, two photodiodes and driving monitoring circuits composed to 12 MESFET's (20) or a transmitter with 4:1 multiplexer circuit (21).

In the InP system, the realizations are more modest but as an example a recent experiment led to the interesting integration of 4 DFB type lasers together with a carrier injection switch (22) to investigate lasing, amplification, detection and switching on the same chip.

3. CONCLUSION

Since the beginning of Integrated Optics two main paths were considered : the hybrid and the monolithic approach.

To day $LiNbO_3$ based circuits are matured and begin to be commercially available in small quantities from several companies. The recent progresses in semiconductor based components will lead in the future to the real monolithic integration of optical and electronic circuitry to achieve complex OEIC chips suitable for applications in fields as different as telecommunication, optical sensors and signal processing.

REFERENCES

|1| Hammer J.M. and Phillips W., Appl. Phys. Letters 24, 11 (1974)

|2| Schmidt R.V. and Kaminow I.P., Appl. Phys. Letters 25, 458 (1974)

|3| Papuchon M. et al, Appl. Phys. Letters 27, 289 (1975)

|4| Gee C.M. et al, Appl. Phys. Letters 43, 998 (1983)

|5| Alferness R.C., Korotky S.K. and Marcatilli E.A., IEEE J. Quantum Electron. QE-20, 301 (1984)

|6| Erasme D. and Wilson M.G.F., Opt. Quantum Electron. 18, 203 (1986)

|7| Nazarrathy M. et al, Paper TuQ37, OFC/IOOC'87, Reno (1987)

|8| Granestrand P. et al, Electron Lett. 22, 816 (1986)

|9| Komatsuk et al, OFC/IOOC'87, 1987, Reno

|10| Link R.A. et al, Electron. Lett. 22, 301 (1986)

|11| Lefevre H.C. et al, Agard Conference on Guided Optical Structures in the Military Environment, Turkey (1985)

|12| Inove H. et al, J. Lightwave Techn. LT3, 1270 (1985)

|13| Lin S.H. et al, Electron. Lett., 21, 597 (1985)

|14| Kapon E. and Bhat R., Appl. Phys. Letters, 50, 1628 (1987)

|15| Koren U. et al, Appl. Phys. Letters, 49, 1602 (1986)

|16| Ishida K. et al, Appl. Phys. Letters, 50, 141 (1987)

|17| Koren U. et al, Appl. Phys. Letters, 50, 368 (1987)

|18| Wood T.H. et al, Electron. Letters 21, 693 (1985)

|19| Wada O., IEEE J. Quant. Electron. QE22, 805 (1986)

|20| Matsueda H. et al, Proc. 12th Int. Symp. Gallium Arsenide and Related Compounds, Kakmiza, Japan (1985)

|21| Carney J.K, GaAs IC Symp. Phoenix, USA (1983)

|22| Inove H. et al, Optoelectronics Dev. and Techn. 1, 137, 1986

Session A4.1

Integrated Optoelectronics I

Chairman: P. Rocchi

Thursday, September 17, 1987

CURRENT-INJECTION ANALYSIS OF INVERTIBLE InGaAsP/InP DOUBLE-HETERO-STRUCTURE BIPOLAR TRANSISTORS

H.G. Bach, N. Grote and F. Fiedler

Heinrich-Hertz-Institut für Nachrichtentechnik Berlin GmbH
Integrierte Optik
Einsteinufer 37, D-1000 Berlin 10, FRG

High-gain invertible double-heterostructure bipolar transistors (DHBT) aimed at laser driver applications have been fabricated on the InGaAsP/InP material system. Forward and reverse current gain evaluation was supported by analyses of forward I-V curves of test-diodes representing the injecting heterojunctions and a current blocking InP homojunction diode. The influence of different LPE methods (step cooling , two-phase solution technique) on the static performance of the DHBTs was studied.

1. INTRODUCTION

For the realization of integrated optoelectronic transmitters for use in long-wavelength (λ=1.3/1.55μm) optical communication systems the application of InP-based bipolar transistors is promising due to their high current handling capability as well as their high switching-speed potential. The employment of invertible double-heterojunction bipolar transistors (DHBT) can facilitate the implementation of the laser driving circuit involved which in a commonly used configuration can be formed by a differential amplifier with a common-emitter current source. In this study invertible InGaAsP/InP DHBTs were characterized with special emphasis laid on the dc-analysis of the forward and reverse mode current behaviour.

2. SAMPLE FABRICATION

Figure 1 shows the structure of the investigated DHBTs which were grown by LPE. The emitter and collector layers (InP:Sn) were symmetrically doped to n=3-4x10^{17}cm^{-3}, whereas an equal or lower doping (Zn) concentration (p=2-4x10^{17}cm^{-3}) was chosen for the quaternary base (0.4μm thick) in order to avoid the risk of detrimental pn-junction displacement due to acceptor out-diffusion. A deep Zn-diffused region underneath the base contact not only facilitates the contacting of the base, but more importantly is to suppress current injection into the extrinsic base in the inverted mode. Non-alloyed ohmic Ti-Au contacts [1] were applied to contact the emitter and collector utilizing a via-hole-technique, whereas alloyed Au/Zn was used for the base contact.

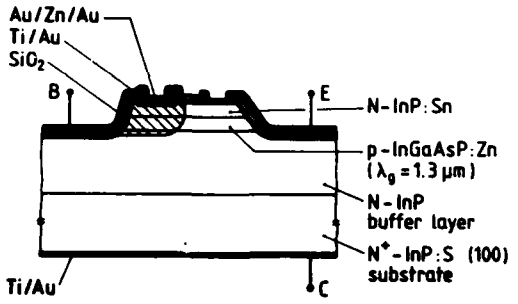


FIGURE 1

Structure of investigated DHBT.

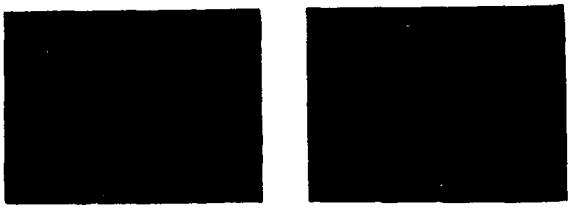
3. RESULTS AND DISCUSSION

Figure 2 shows typical forward (a) and reverse (b) common-emitter characteristics (note: forward mode=emitter up). Appreciably high current gain values of ≈5000 and ≈300 (at I_C>10mA) were achieved in the respective modes of operation.

To get a deeper insight into the current behaviour the related Gummel-plots were analysed. Figure 3 depicts the respective I-V_{be} curves (V_{bc}=0) for the

This work was conducted under the ESPRIT programme (project 263 B).

emitter-up configuration (full lines) and the inverted transistor (dashed lines).



a) forward mode b) reverse mode

FIGURE 2

Common-emitter characteristics of DHBT with doping concentrations of $n=4x10^{17}cm^{-3}$ (emitter/collector) and $p=2x10^{17}cm^{-3}$ (base).

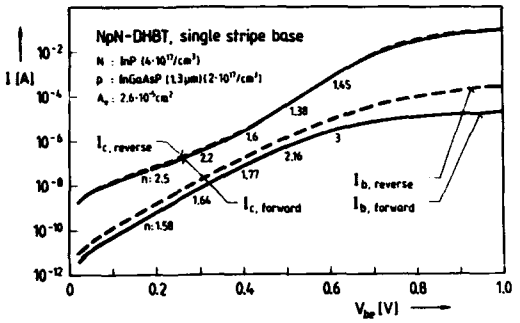


FIGURE 3

Gummel-plot of DHBT shown in figure 2.

The Gummel-plot is seen to be well-behaved in the intermediate collector current range of $\approx 10^{-4}$A where the ideality factors of the collector current curves approach $n \approx 1.3$. In the high current regime the current curves tend to saturate which is well understood to be due to the influence of the series resistances of the emitter and the base layer rather than to a drastic increase of the ideality factor.

At low currents surprisingly high current gain values of up to roughly 10^3 are suggested from the measured curves. However, these values are considerably higher than those obtained from the related output characteristics displayed in figure 4. This discrepancy can be explained if a leakage current is assumed to flow between emitter and collector, possibly via an n-type surface channel across the edge of the p-doped base layer. This interference can be eliminated by extrapolating the medium range (10^{-4}A) Ic-curves towards the lower currents resulting in good agreement of the gain values determined in the two different ways. For the particular device shown in figure 4 the gain amounts to about 50 at $I_C<100$nA which is yet remarkably high and compares reasonably well with results reported by Nottenburg et al. for GSMBE grown step-graded DHBTs [2].

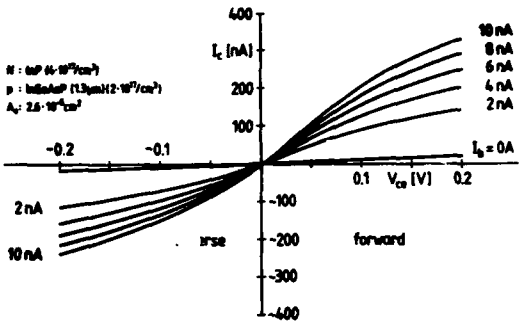


FIGURE 4

Low current output characteristics of invertible DHBT.

Such a gain behaviour which, to our knowledge, has not been observed in GaAlAs HBTs render these transistors attractive also for use in optoelectronic receivers.

Another note-worthy feature is the good symmetry of forward and reverse gain in the low current regime which is in contrast to the behaviour at high currents. This is illustrated in figure 5 which shows the common-

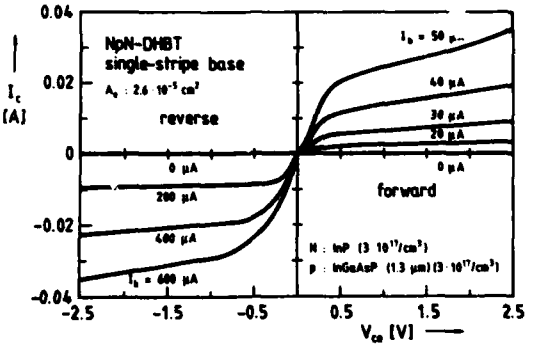


FIGURE 5

DHBT output characteristics at high collector currents (doping concentrations = $3x10^{17}cm^{-3}$ for emitter, base and collector).

emitter characteristics of a DHBT (with slightly different doping concentrations) for either operating mode in the mA-collector current range. The output characteristics exhibit a satisfactory saturation behaviour in both operating modes. The current level of some 50mA will be adequate for laser drive applications. Current gain values can be seen to be ≈600 and ≈60 in the forward and the inverted mode, respectively, giving a forward-to-reverse gain ratio of one order of magnitude.

To analyze this discrepancy the electrical properties of the different diodes incorporated in the DHBT were characterized in more detail utilizing large-area test-diodes to represent the heterojunctions and the diffused homojunction diode. Figure 6 shows the average forward I-V curves for the lower (i.e. base/collector) heterodiode and the diffused homodiode taken from a lot of some 10 samples each.

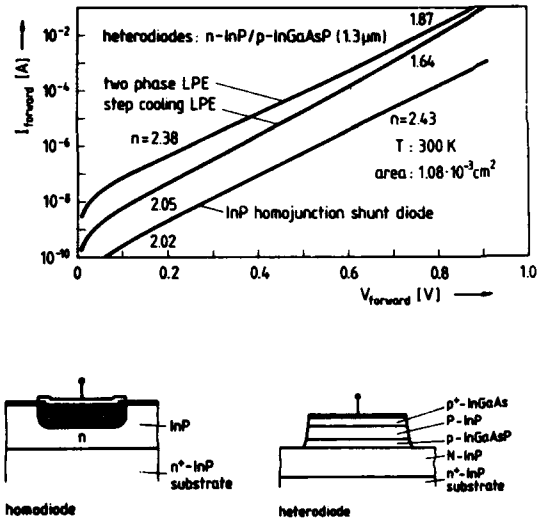


FIGURE 6

Forward current I-V curves of large-area test-diodes as shown in the inset. In case of the heterodiodes step cooling LPE as well as the "two-phase-solution" technique were employed for the growth of the quaternary layer for comparison.

Using modelling parameters derived from these I-V curves (see table 1) and taking into account the lateral base resistance the effect of current injection into the extrinsic base region on the current gain was calculated. As a result it can be concluded that in the current range of interest (≤100mA) this parasitic current is sufficiently suppressed so as not to affect the current gain significantly.

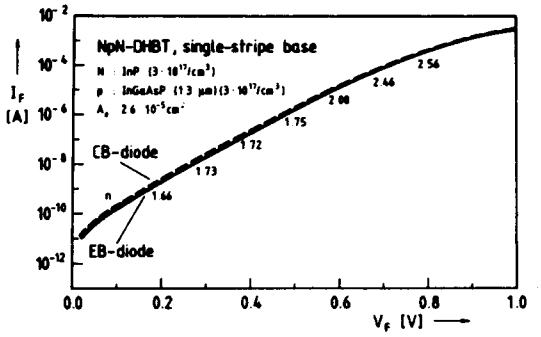


FIGURE 7

Forward current I-V curves of individual emitter and collector diodes of DHBT presented in figure 5.

As a further cause for the high-current gain asymmetry, it was suspected that both heterojunctions of the DHBT exhibit different carrier injection properties at higher current levels. Such differences should be reflected in the forward current/voltage characteristics of the individual emitter/base and base/collector diodes. Respective I-V curves, measured on the particular transistor presented in figure 5, are plotted in figure 7. The curves can be seen to almost completely coincide over nearly 10 decades of current proving equivalent saturation currents and ideality factors. Hence, these measurements actually do not indicate the presumed difference in the injection behaviour leaving the above problem unsolved. However, the coincidence of the current curves again confirms the effectiveness of the InP homojunction diode to prevent parasitic base current flow.

Moreover, test-diodes and DHBTs were made the quaternary layers (base) of which were grown by the "two-phase-solution" instead of the "step cooling"-LPE technique in order to study the effect on current gain. The forward I-V curve of the test structure is also

Table 1

Ideality factors and saturation currents of diffused InP homojunction and heterojunction diodes grown by different LPE techniques (test-diodes) (Q:InGaAsP(1.3μm)).

test sample	p doping [cm^{-3}]	n doping [cm^{-3}]	low current Js[A/cm^2]	high current Js[A/cm^2]	low current ideality fac.	high current ideality fac.
diffused InP homojunction diode	p^+	$2.5*10^{17}$	$4.4*10^{-8}$	10^{-6}	2.02	2.43
lower heterodiode p-Q/n-InP (Q-layer: step cooling LPE)	$1.5*10^{17}$	$4*10^{17}$	$8.8*10^{-7}$	$3.2*10^{-8}$	2.05	1.64
lower heterodiode p-Q/n-InP (Q-layer: "two-phase-solution" LPE)	$2.4*10^{17}$	$4*10^{17}$	$1.5*10^{-7}$	$7.7*10^{-7}$	2.38	1.87

depicted in figure 6 indicating higher saturation currents and higher ideality factors as compared to diodes grown by "step cooling" LPE. These measurements suggest inferior gain properties which were preliminarily verified by the gain figures of "two-phase-solution" DHBTs being typically lower by roughly one order of magnitude. However, this might also be due to an additional reduction of the base transport factor.

4. CONCLUSIONS

High gain invertible double-heterojunction bipolar transistors have been fabricated and analyzed. Step cooling LPE proved to be an easy method to obtain transistors with sufficiently symmetric operation which will greatly simplify further integration of a laser driver circuit.

ACKNOWLEDGEMENTS

The authors wish to thank H. Schroeter-Janßen, M. Mahnkopf, I. Tiedke and B. Lehmann for their expert technical assistance during the course of this work.

REFERENCES

[1] Bach, H.G., Grote, N. and Niggebrügge, U.,Advanced Materials for Telecommunications 1986, XIII,Les Éditions des Physique (1986) edited by Glasow, P.A. et al., pp. 461-466.

[2] Nottenburg, R.N., Panish, M. B. and Temkin, H., Inst. Phys. Conf. Ser. No. 83 (1987) 483.

HETEROJUNCTION BIPOLAR PHOTO-TRANSISTORS FOR HIGH SPEED LOGIC AND COMMUNICATION APPLICATIONS

A.J. Doherty and W.S. Truscott

Joint Laboratory of Physics and Electrical Engineering
Department of Electrical Engineering and Electronics
University of Manchester Institute of Science and Technology
P O Box 88, Manchester M60 1QD, U K

We report the results of an experimental study on a GaAs/AlGaAs double heterojunction bipolar transistor structure which can be integrated with fast optical detectors. Results are presented showing that the structure can be fabricated into transistors with f_ts over 100 GHz, and that the photodetecting structure has a good sensitivity with an incident wavelength of 830 nm.

1.INTRODUCTION

Heterojunction bipolar transistors are recognised as devices suitable for the very highest speeds. In communications they will be used at the opto-electronic interface in fibre optic systems where the data rate is extremely high. In logic applications they will be used in specialised circuits requiring the highest possible clock rate. We have been studying a structure which allows optical waveguides to be integrated with heterojunction bipolar transistor ics. Such circuits will be useful for both communications and logic applications; in the latter an optically propagated clock signal would give a very fast rise time and a high degree of synchronism.

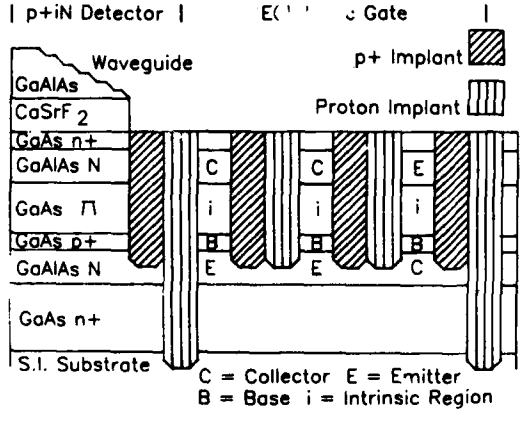


FIGURE 1

The proposed integrated phototransistor and waveguide structure.

The structure that is the subject of this study is shown in Figure 1. The optical waveguide at the top left of the diagram consists of AlGaAs grown epitaxially onto $Ca_xSr_{1-x}F_2$. The composition of the $Ca_xSr_{1-x}F_2$ can be chosen so that it lattice matches the underlying GaAs. The light is diverted into the p^+iN photodetector using an angled edge as a mirror or by a grating structure. The photo-detector, which is the lower structure on the left of the diagram, is a GaAs/AlGaAs p^+iN photodiode. Contact to the top n-type layer can be via metallisation at the side of the waveguide normal to the plane of the paper. Contact to the p^+ layer is achieved using ion-implantation of a suitable dopant, e.g. Be, Mg. Isolation is by ion implantation of protons. The logic gate, which is the structure in the centre and the right of the diagram, is an ECL gate which has two collector up switching transistors fed by an emitter up current source transistor. The intrinsic region of the p^+iN detector forms part of the collector/base junction of the collector up transistors. The presence of this region increases the transit time of electrons across the depletion region, but as shown later, does not necessarily decrease the f_t of the transistor since there is a related decrease in the collector capacitance. The intrinsic region in the base/emitter junction of the emitter up transistor will reduce the injection efficiency since more recombination will occur in the depletion region. Contact to the lower layers is not necessary because of the logic family chosen. Ion implantation is used for both isolation and the base contacts. Kroemer [1] has proposed, as an electronic logic gate, a related structure

without an intrinsic layer. The transistor structure includes a p^+iN optical detector, though we do not expect it to be used as a conventional photo-transistor since p^+iN diode detectors have faster response speeds than phototransistors. In addition, extra layers of AlGaAs can be grown to act as light emitting diodes, though this is not shown on the diagram.

2. EXPERIMENTAL

As a preliminary investigation into the above structure we have grown, fabricated and tested the structures shown in Table 1. These layers were grown by MBE on SI substrates, the order of growth was that shown in the table. Although the structures are designed to be fabricated using ion-implantation to contact the base and isolate the devices, the results presented have been obtained using mesa etching for these tasks.

A mask set giving 11 different geometries, all rectangular in shape has been employed which allows us to determine geometry dependent parameters. A section of the mask is shown in Figure 2. We have measured the optical and electrical DC characteristics of these devices including the measurement of the photo-current as a function of the wavelength of the incident light. The light was shone onto the devices through the optics of the probing station which was fed via an optical fibre with the output of a monochromator.

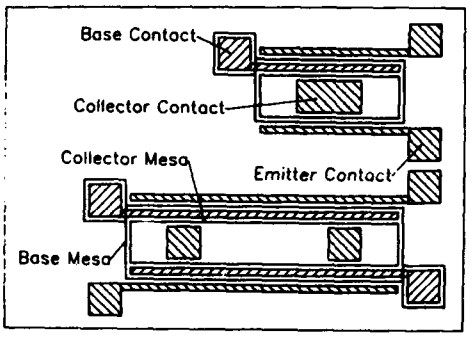


FIGURE 2

A section of the mask set showing two different geometries.

3. RESULTS

The transistors were contacted using Au/AuGe evaporated onto n-type GaAs and Au/Zn/Au evaporated onto p-type GaAs. Typical values for the specific resistance of these contacts were found to be $4 \times 10^{-4} \Omega\ cm^2$ and $4 \times 10^{-3} \Omega\ cm^2$ respectively. The transistors exhibited good Ic/Vce transfer characteristics with current gains around unity. These gains are low because the device geometry is such that up to half the emitter current is injected directly under the base contact. The gains are expected to rise when ion implantation is used to equalise the *active* emitter and collector areas [2].

For a device with a collector area of 0.11 mm^2 and emitter area of 0.19 mm^2 the measured base, emitter and collector resistances were 370 Ω, 70 Ω and 330 Ω respectively. The emitter/base junction capacitance, measured at 1 volt forward bias, was 470 pF, while for a 3000 Å intrinsic region in the collector base junction the capacitance was 55 pF at −1 V reverse bias. The collector capacitance varied very little with reverse bias because of the presence of the intrinsic layer. The reverse biased saturation current of the collector/base junction was less than 10^{-9} Amps up to breakdown at −16 volts.

A measure of the performance of a transistor is given by its f_t which can be calculated using the equation:

$$\frac{1}{2\pi f_t} = r_e C_{TE} + \frac{W_b^2}{2.43\ D_b} + \frac{X_d}{2V_{SL}} + (r_e + R_E + R_C) C_{TC}$$

where
- r_e = a.c. resistance of the emitter/base junction
- R_E, R_C = emitter and collector resistance
- C_{TE} = emitter/base junction capacitance
- C_{TC} = collector/base junction capacitance
- W_b = base width
- D_b = diffusion co-efficient of electrons in the base

x_d - collector/base junction depletion width

V_{SL} - saturation velocity of electrons in collector/base junction

The last two terms in eqn 1 are dependent upon the intrinsic layer thickness. Increasing the intrinsic layer thickness increases the third term which represents the electron transit time across the base/collector depletion region but reduces the last term which represents the base/collector capacitance charging time by reducing the capacitance.

MATERIAL	TYPE	DOPING	THICKNESS
GaAs	n+	1e18 /cc	3500 A
GaAlAs	N	5e17 /cc	1500 A
GaAs	p+	1e19 /cc	750 A
GaAs	Π	undoped	2000 A
GaAlAs	N	5e17 /cc	1500 A
GaAs	n+	1e18 /cc	200 A

Table 1 A list of the layers grown on S.I. substrates

We have used the above measured values, together with previously reported specific contact resistances of 1×10^{-6} and 5×10^{-7} Ω cm² for n and p type contacts [3], to calculate parameters for a 2 x 20 μm transistor fabricated with 1 μm line widths and ion implantation for base contacts. We find that the scaled parameters are: R_b = 97 Ω, R_E = 20 Ω, R_C = 5 Ω, C_{TC} = 12.2 fF for a 3000 Å intrinsic layer, and C_{TE} = 84 fF. Figure 3 shows a plot of f_t versus depletion layer thickness obtained for a current density of 5×10^3 A cm^{-2} and a saturated velocity of 2×10^7cm^{-1}. This value of the saturated velocity may be lower than that achieved in practice because of ballistic transport across the thin intrinsic region. The value of f_t has a maximum of 107 GHz at an intrinsic region width of 2000 Å. This calculation shows that the inclusion of the intrinsic region can improve the performance of the device. This effect was first demonstrated by Early [4] in 1952. The value of the intrinsic region thickness for the maximum f_t is dependent upon the structure used. If the values of R_E, R_C, r_e and C_{TC} are increased the intrinsic region thickness to give the maximum f_t is also increased.

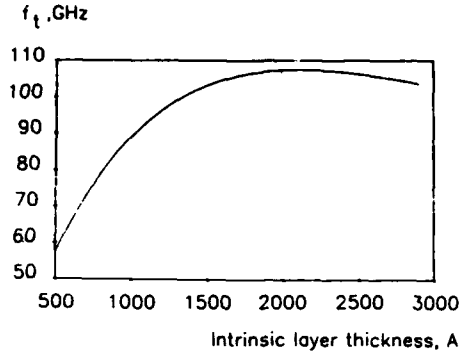


FIGURE 3

A plot of f_t against layer thickness

Measurements of photo-sensitivity have been made on three different structures. Figure 4 shows the optical response, which is proportional to AW^{-1} for a given device geometry, plotted against the wavelength of incident light. The devices corresponding to curves 1 and 3 have larger areas than that of curve 4 which in turn has a larger area than those of curves 2 and 5. The sensitivity curves reflect these different areas. As the energy of the incident light is increased a significant photo-current occurs just before the band edge of the intrinsic GaAs. This may be accounted for by the bandwidth of the incident light, which we estimated to be approximately 20 nm. At a wavelength of 830 nm the sensitivity for an intrinisic region thickness of 2000 Å is 94% that of a 3000 Å thickness and the sensitivity for a thickness of 1000 Å is only 73% of the 3000 Å thick intrinsic region. This can be explained by the exponential manner in which the light is absorbed in the intrinsic region. Thus a 2000 Å layer thickness gives a good optical sensitivity and, as seen earlier for the scaled device geometry, gives a maximum value for the f_t. The sensitivity increases past the energies of the bandgap of the AlGaAs, we believe that this is due to a current of holes generated with the AlGaAs collector diffusing to the intrinsic region and being swept across by the electric field. This wavelength region is not suitable for device operation because the response speed would be slower than at longer wavelengths where drift currents dominate.

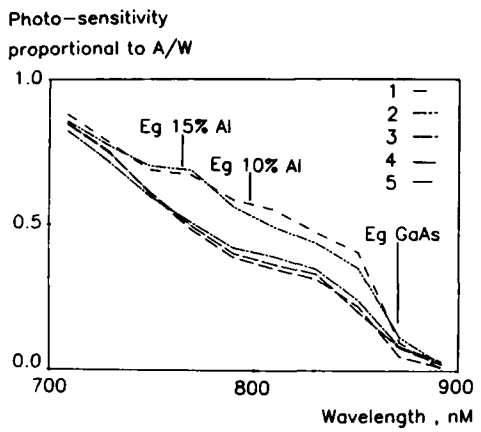


FIGURE 4

A plot of photo-sensitivity which is proportional to A/W against the wavelength of incident light for different structures. The device represented by curve 1 has 10% Al in the collector and an intrinsic region of 3000A, the device represented by curve 2 has again 10% Al and an intrinsic region of 2000A, while curves 3, 4, 5, represent devices with 15% Al and an intrinsic region of 1000A, but with the area of the devices decreasing respectively. Each curve has been normalized by the effective area of the device.

4. CONCLUSION

We have proposed a structure which combines optical waveguides, optical detectors and fast switching transistors, where the base collector junction of the transistor acts as a p^+iN photodetector. The devices we have grown and fabricated show good dc transistor characteristics. Using the measured electrical parameters of the device we have shown that a device with 2 x 20 μm geometry will have a maximum f_t of 107 GHz with an intrinsic region width of 2000 Å. The sensitivity of the p^+iN photo-detector for 2000 Å intrinsic layer was 94% that of a 3000 Å layer at a wavelength of 830 nm. Thus the structure with a 2000 Å thick intrinsic region can act as an optimised fast transistor and a sensitive fast optical detector. The optimum value for the intrinsic layer thickness is determined by the device structure and the wavelength of the light to be used. This technology can be applied to other material systems, such as GaInAs/InP structures for long wavelength applications.

ACKNOWLEDGEMENTS

This work has been supported by the SERC through a number of research grants and a Total Technology Award which is in conjunction with STC Technology Ltd., Harlow, UK.

REFERENCES

[1] H. Kroemer, "Heterostructure Bipolar Transistors and Integrated Circuits", Proc. IEEE, Vol. 70, p.13, 1982.

[2] S. Adachi and T. Ishibashi "Collector-up HBT's Fabricated by Be^+ and O^+ Ion Implantations", IEEE Electron Device Letters, Vol.EDL-7, No.1, January 1986.

[3] B. Kim, H.Q. Tseng, S.K. Tiku and H.P. Smith "AlGaAs/GaAs Heterojunction Power Transistors", Electron. Lett., Vol. 21, p.258, 1985.

[4] J.M. Early "Effects of Space-charge Layer Widening in Junction Transistors", Proc. IRE, 40, p.1401, 1952.

AlInAs/GaInAs MESFETs GROWN BY MBE

L. Giraudet, M. Allovon, J.Ch. Renaud, A. Scavennec

Centre National d'Etudes des Télécommunications
Laboratoire de Bagneux
196, Avenue Henri Ravera - 92220 Bagneux - FRANCE

ABSTRACT : Field Effect Transistors have been fabricated on AlInAs/GaInAs/InP heterostructures grown by molecular beam epitaxy (MBE). Very low gate leakage currents (20 nA at - 5 V gate bias) were obtained together with transconductances of 90-100 mS/mm on devices with a 3 μm gate length. High frequency measurements were performed showing a cutoff frequency for the maximum available gain of 20 GHz. These characteristics make this type of FET very promising for microwave and integrated optoelectronics applications such as large bandwidth photoreceivers.

1. INTRODUCTION

Optoelectronic integration on InP substrates for the 1.3 - 1.6 μm wavelength range is expected to lead to large bandwidth, high performance, compact circuits with good reproducibility. Considerable effort has been devoted worldwide, to the development of a GaInAs FET technology in order to take advantage of its high channel velocity (2-3.10^7 cm/s). Presently various structures have been investigated, including J-FET, MISFET and MESFET (Schottky gate FET).

The last one seems very promising [1] : it is not prone to current drift phenomena (as MISFET) and very short gates can be easily achieved. In this paper, the realization and characterization of both Schottky diodes and high transconductance MESFETs (about 100 mS/mm for L_G = 3 μm), using the AlInAs/GaInAs system, are reported.

The behaviour of such heterojunction-based devices is strongly dependent on the material quality. Several advantages such as : uniformity of doping levels and layer thicknesses, interface abruptness, variable AlGaInAs composition, make the Molecular Beam Epitaxy a good candidate for this objective.

2. SCHOTTKY GATE FET FABRICATION

The FET multilayer structure, grown by MBE on Fe doped, semi-insulating InP substrate, is the following (see fig. 1) :

-0.3 μm AlGaInAs buffer layer,
-0.3 μm Si-doped, n-type GaInAs channel layer,
-0.2 μm undoped AlInAs barrier layer.

Hall test patterns included in the set of masks used for the transistors processing allow measurement of the channel doping level (3.3 x 10^{16} cm^{-3}) and Hall mobility (9200 cm^2/Vs at 300 K). Such a high mobility could not be achieved without the AlGaInAs buffer layer. This high mobility demonstrates the good quality of the MBE material. The resulting channel sheet resistance is about 700 Ω/□ .

A cross sectionnal view of the transistor is shown on figure 1. The devices are first mesa isolated with a H_3PO_4 : H_2O_2 : H_2O etch down to the SI-InP substrate. A recess is then patterned into the AlInAs layer in order to evaporate source and drain ohmic contacts. A thin AlInAs layer is left above the GaInAs channel in the source and drain contact regions. This is expected to ease contacting an

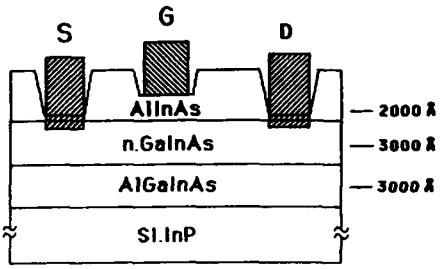


FIGURE 1

Schematic cross-section of the AlInAs/GaInAs transistor.

inversion layer formed at positive gate bias at the GaInAs/AlInAs interface. Contact resistivity of 10^{-5} Ω cm² have been measured after alloying of the AuGeNi contacts. The last step consists in evaporating the TiAu gate on the AlInAs layer, preliminarily recessed down to 1000 A.

3. SCHOTTKY DIODE

In order to obtain a good barrier height and low reverse current, a wide gap material is required on top of the active InGaAs layer: as described above, undoped AlInAs, grown at high temperature, has been chosen. Forward and reverse I(V) characteristics of the Schottky gate are presented in figure 2. At - 5 V gate bias, a very low leakage current close to 10^{-3} A/cm² is observed.

Measurements of the temperature dependence conducted at low reverse gate voltage corresponding to usual transistor operating conditions, indicate that generation-recombination in the wide gap material is the main component of the Schottky diode reverse current. Because of the high resistivity of the AlInAs layer, the ideality factor n of the diode could not be determined. Nevertheless, the Schottky barrier height on AlInAs, not well known at the moment but probably lying between 0,58 eV [2] and 0,8 eV [3], appears to be consistently much larger than it is on GaInAs (0,2 eV) and is suitable for a MESFET structure. The very low reverse currents measured are especially interesting in the objective of the monolithic integration of a PINFET photoreceiver with very low noise level.

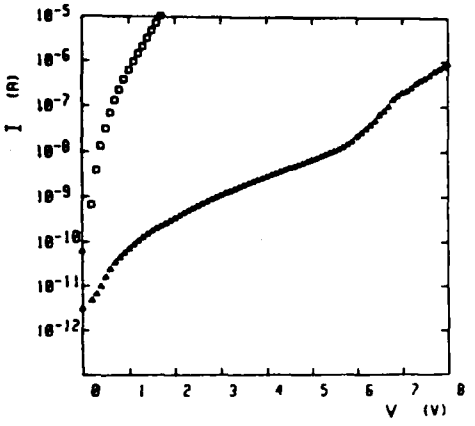


FIGURE 2

Forward (□) and reverse (△) current of 3 x 300 μm TiAu/AlInAs/GaInAs diode.

4. TRANSISTOR CHARACTERISTICS

Typical DC characteristics of the FETs are shown in figure 3. Transconductances of 90-100 mS/mm are obtained on devices with 3 μm gate length, 300 μm gate width, and a drain to source spacing of 5 μm. The saturation drain current is close to 60 mA. Characterization using classical MESFET models [4] indicates a channel sheet resistance within 10 % of the preliminary Hall measurements together with a velocity saturation type of behaviour even for a gate length larger than 3 μm. Figure 4 shows the FETs transconductance versus gate voltage. The transconductance is maximum for a gate bias very close to V_G = 0 V and decreases with reverse gate bias with a standard linear behaviour. For forward gate bias, g_m decreases quite rapidly suggesting a parallel conduction in the AlInAs barrier

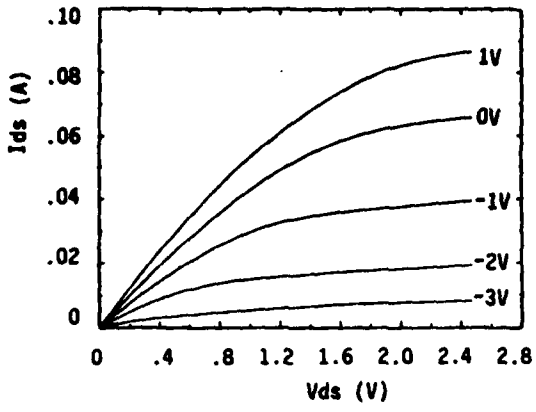


FIGURE 3

DC drain current-voltage characteristics of a AlInAs/GaInAs MESFET. Lg = 3 μm, Wg = 300 μm.

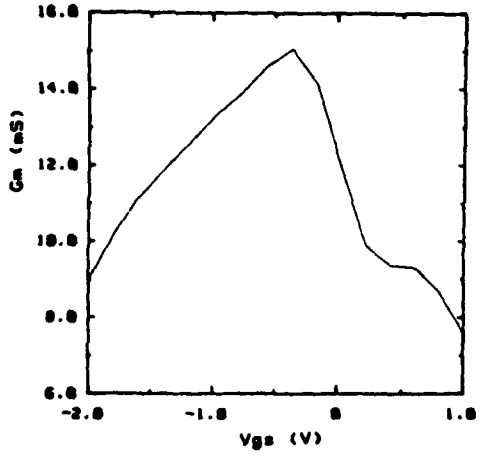


FIGURE 4

Typical g_m versus gate voltage characteristics (W_g = 150 μm).

layer as it is observed usually in TEGFETs [5]. Nevertheless, the transconductance of the AlInAs/GaInAs MESFETs remains higher than $g_{m\ max}/2$ over quite a large gate voltage range, typically beetween + 1 V and - 2 V.

5. HIGH FREQUENCY MEASUREMENTS

After mounting on alumina substrates, high frequency measurements have been performed in the 1 GHz - 18 GHz range. The current cutoff frequency is about 10 GHz, which is in good agreement with values of C_{GS} (~ 0,5 pF) and g_m (~ 30 mS) measured at low frequency. Figure 5 shows the variations of the maximum available gain (MAG) as computed from S parameters. The slope of the curve (- 16 dB/dec.) is close to the expected -20 dB/dec, and a MAG cutoff frequency of about 20 GHz can be inferred. These values, obtained for a gate length of 3 μm, confirm the attractive potentialities of the devices at high frequency.

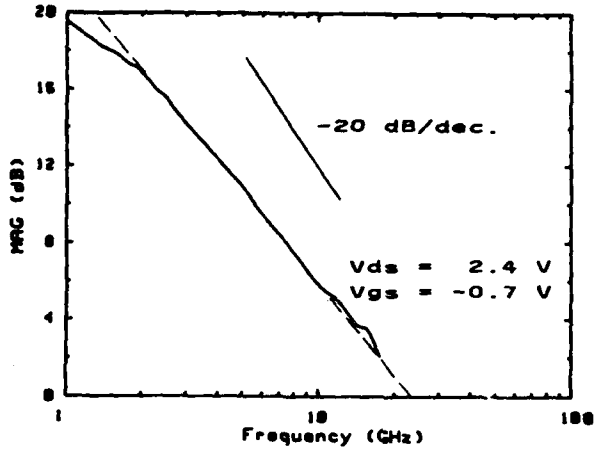


FIGURE 5

Maximum available gain versus frequency characteristic.

6. CONCLUSION

AlInAs/GaInAs MESFETs with very low gate leakage current, good transconductance and high cutoff frequency have been successfully fabricated. These characteristics are very attractive for optoelectronics integrated circuits such as PIN-FET photoreceivers.

Moreover, by reducing the FET gate length and channel sheet resistivity (respectively 3 μm and 700 Ω/□ presently), higher transconductances may be easily achieved, and higher transition frequency obtained.

ACKNOWLEDGMENTS

The authors would like to thank J. Dangla, L. Nguyen and M. Laporte for assistance in DC and HF measurement and characterization.

REFERENCES

[1] M.D. Scott, A.H. Moore, I. Griffith, R.J.M. Griffiths, R.S. Sussmann and C. Oxley, "AlInAs/GaInAs MESFETs by MOCVD", Inst. Phys. Conf. Ser. No. 79 : Chapter 9. Int. Symp. GaAs and Related Compounds, Karuizawa, Japan, 1985.

[2] H. Ohmo, C.E.C. Wood, L. Rathbun, D.V. Morgan, G.W. Wicks, and L.F. Eastman, J. Appl. Phys. 52, 4033 (1981).

[3] S. Hata, H. Yasaka, K. Nakashima, H. Asahi and S. Uehara, "Monolithic integration of a Schottky photodiode and a MESFET using an InAlAs/InGaAs heterostructure", 12th European Conference on Optical Communication, Barcelona 1986, vol I.

[4] D. Boccon-Gibod, "Modèle analytique et schéma équivalent du transistor à effet de champ en arseniure de gallium", Acta Electronica. 23. (80). p. 99.

[5] L.F. Eastman, "III-V heterojunction field effect transistor using Indium alloys", Proceedings IEDM, Los Angeles, December 1986.

Session B4.1

Bipolar Modelling I

Chairman: H.C. De Graaff

Thursday, September 17, 1987

THE INFLUENCE OF EMITTER SIDEWALL INJECTION ON TRANSISTOR NOISE FIGURE

H.R. Claessen, J.A.M. Geelen, H.C. de Graaff
Philips Research Laboratories
P.O. Box 80.000
5600 JA Eindhoven - The Netherlands

1. INTRODUCTION

High speed bipolar transistors are characterized by shallow junctions and small lateral dimensions. As a consequence such devices can suffer considerably from parasitic effects of the junction sidewalls, especially when the base-emitter junction is considered. One major effect is the direct injection of base current into the emitter sidewall. Hurkx [1] has shown by 2-D device simulations that for emitter junction depths between 0.1 and 0.3 μm , and emitter widths varying from 0.3 to 4.0 μm ,10 to 75 % of the base current may be injected directly into the emitter sidewall, depending on the ratio between junction depth and emitter width. This effect does not only influence the current gain, but it also causes an important increase in transistor noise figure.

It will be shown that the transistor noise figure for the white spectrum as calculated by van der Ziel and Sze [2,3] , based on the equivalent-circuit diagram given in fig. 1, does not apply to bipolar devices exhibiting the aforementioned sidewall effect. A comparison of measurements and calculations reveals that noise figure calculations based on the commonly used model as stated in fig. 1 are too low, especially at high current levels. This effect also has its implications on transistor base resistance measurements using the low frequency thermal-noise method. As discussed by Unwin and Knott [4] , this method has been found to yield accurate results for values of the total base resistance R_b down to 5 Ω . However, it will be shown that by neglecting sidewall injection for typical HF transistors with narrow emitters, extremely large errors in the calculation of R_b may occur.

2. SIDEWALL INJECTION INTO THE EMITTER

Each vertical bipolar transistor will show in some extend parasitic electrical effects of the junction sidewalls. In this section we will discuss the phenomenon of base current injection into the sidewall of the emitter-base junction.

The fraction of the DC base current I_b which is injected into the sidewall of the junction is denoted by α , see fig. 2. Theoretically α depends on the shape of the emitter doping profile, the bulk and surface (contact) recombination rate in the emitter, and the distance between emitter contact and junction sidewall (indicated by d_1 in fig. 2), see also [1] .

In practice α can adequately be determined by separating the bottom- and sidewall components of the base current, using several different emitter geometries. For two different HF bipolar processes (process A and process B), we determined the specific sidewall- and bottom components of the ideal base current by measuring transistors with different geometries. From these specific components it is possible to calculate α as a function of emitter width, see fig. 3.

Process A (x_{je} = 0.26 μm ; x_{jb} = 0.44 μm ; F_t = 10 GHz) and process B (x_{je} = 0.35 μm ; x_{jb} = 0.63 μm ; F_t = 7 GHz) do not only differ in emitter and base junction depth (x_{je} and x_{jb}), but process A also has a shallow (0.1 μm) oxide isolation layer at the edge of the e-b junction.

It can be observed in fig. 3 that process A shows less sidewall injection than process B as was expected due to the shallower emitter and the extra oxide isolation stripe at the e-b junction edge.

From Hurkx's calculations [1] it follows that most of the decrease in α should be attributed to the decrease in emitter junction depth from 0.35 μm to 0.26 μm and not to the influence of the oxide isolation layer.

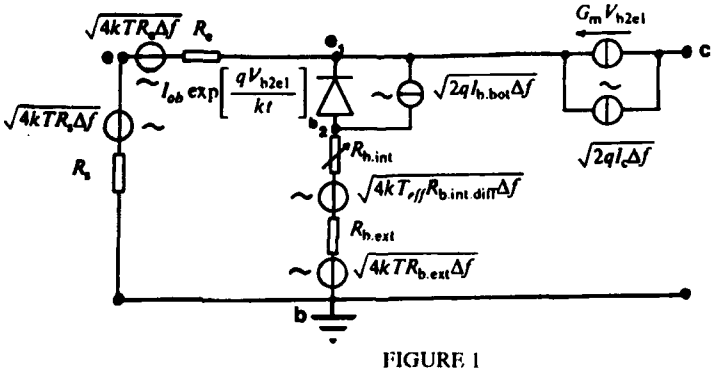


FIGURE 1

Full equivalent noise circuit of the transistor in common base configuration. T, q, k, and Δ f stand for device temperature, elementary electron charge, Boltzmanns constant, and effective noise bandwidth. Source-, emitter-, intrinsic-, and extrinsic base resistance are indicated by R_s , R_e , $R_{b.int}$, $R_{b.ext}$.
The transconductance is given by G_m . For each resistor R a thermal noise emf $\sqrt{4kTR_b\Delta f}$ is present. The mean-square value of the current generators representing the base-emitter and base-collector shot noise equals $\sqrt{2qI\Delta f}$ where I stands for the DC base- or collector current I_b or I_c respectively.

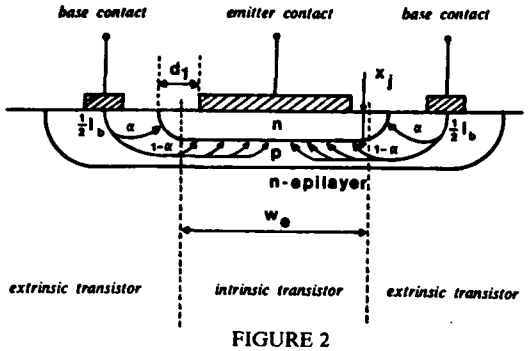


FIGURE 2

Schematical cross-section of bipolar npn transistor showing base current flow.

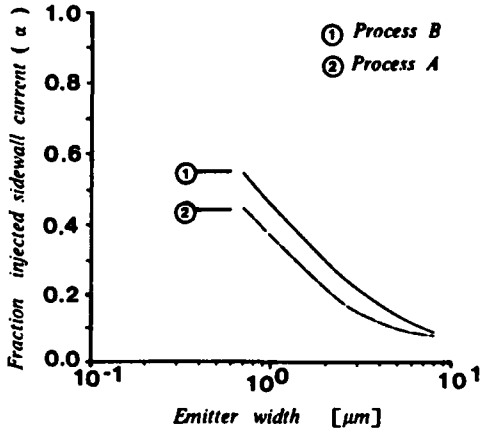


FIGURE 3

Plot of α vs emitter width for process A and B.

3. DERIVATION OF THE NOISE MODEL

In order to derive a noise model it is necessary to describe first the DC behaviour of the base current.

In the previous section α has been defined as the fraction of DC base current which flows directly into the emitter sidewall. However, it must be emphasized that this concept only holds for sufficiently low DC base current levels. For higher DC base currents, the effect of emitter current crowding may occur [5,6]. As a consequence the fraction of DC base current which is injected into the emitter sidewall increases for increasing base current. This effect can adequately be modelled by defining a current depending internal base resistor $R_{b.int}$ and splitting up the e-b junction diode in a bottom- and sidewall component, see fig. 4. The current dependence of $R_{b.int}$ is modelled by the following relation between internal base current $I_{b.bot}$ and internal base voltage drop V_{b1b2}:

$$I_{b.bot} = \frac{2\,V_t}{3R_{b.int}}\left[\exp\left(\frac{V_{b1b2}}{V_t}\right) - 1\right] + \frac{V_{b1b2}}{3R_{b.int}} \qquad (1)$$

with $V_t = \frac{kT}{q}$,

k	Boltzmann's constant	(J/K),
T	Device temperature	(K),
q	Elementary electron charge	(C),
$R_{b.int}$	Low current value of the internal base resistance	(Ω).

Current crowding affects the noise behaviour of the internal base resistor [7]. It is possible to approximate this phenomenon by defining an effective noise temperature for the internal base. This effective temperature depends on the ratio between the low and high current differential resistance of the internal base region. The relation between effective noise temperature T_{eff} and device temperature T is given by eq. (2).

$$T_{eff} = T \times \left[\frac{R_{b.int.diff}\,(high\ current)}{R_{b.int.diff}\,(low\ current)}\right]^{1/4} \qquad (2)$$

with $R_{b.int.diff}$ the differential resistance of $R_{b.int}$ (Ω).
For low currents T_{eff} equals T and $R_{b.int.diff}$ equals $R_{b.int}$

As indicated in fig. 4 the splitting up of the e-b junction diode is done in such a way that for low currents α equals the fraction of the total base current that flows through the sidewall diode.

Charge modulation of the internal base resistance is taken into account according to [8,9] by considering the ratio of fixed base charge and the total charge in the base modulated by the depletion charges and the injected charge of the carriers. Charge modulation has an effect on the value of the internal base resistance, but not on its effective noise temperature. The equation we use to describe the charge modulation is given below:

$$R_{int.mod} = R_{b.int} \times \frac{Q_{b0}}{Q_{b0} + Q_{te} + Q_{tc} + Q_{be}} \qquad (3)$$

with

$R_{int.mod}$	the modulated value of $R_{b.int}$	(Ω),
Q_{b0}	the fixed base charge	(C),
Q_{te}, Q_{tc}	the emitter and collector depletion charges	(C),
Q_{be}	the injected electron charge	(C).

Combining eq. (1-3) with the equivalent circuit-diagram shown in fig. 4 leads to the transistor DC-model.

As in general the Gummel number of the extrinsic base (the base region besides the emitter) will be an order of magnitude higher than the Gummel number of the intrinsic base (underneath the emitter), the collector current in our model is controlled only by the potential difference accross the internal base-emitter junction diode (V_{b2e1}). This means that collector current caused by injection of electrons from the emitter into the extrinsic base region is neglected.

In order to make the equivalent circuit-diagram of fig. 4 suited for noise analyses, we have to perform the following steps: after calculation of $I_{b.bot}$, $I_{b.sw}$, and the collector current I_c from a DC bias condition, both diodes and $R_{b.int}$ have to be replaced by their differential resistance:

sidewall diode $\quad R_{diff.sw} = \frac{kT}{qI_{b.sw}}$,

bottom diode $\quad R_{diff.bot} = \frac{kT}{qI_{b.bot}}$,

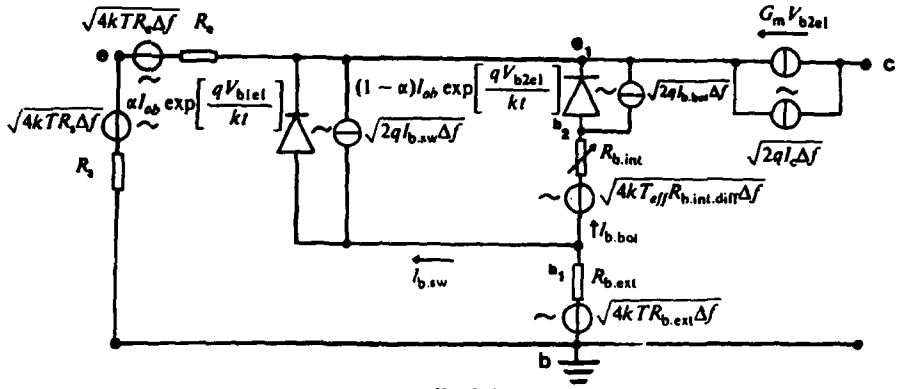


FIGURE 4

Altered circuit-diagram of fig.1 in which the e-b junction has been split up into a sidewall and a bottom diode. Each diode has its own shot noise source. The DC bottom- and sidewall base current component are indicated by $I_{b.bot}$ and $I_{b.sw}$ respectively.

$$\text{internal base resistor} \quad R_{b.int.diff} = \frac{3R_{b.int}}{2\exp\left(\frac{V_{b1b2}}{V_t}\right) + 1}.$$

G_m must be replaced by its small signal value g_m :

$$g_m = \frac{qI_c}{mkT} \quad (A/V),$$

with m the non ideality factor of the collector current [10] .

4. EXPERIMENTS AND MODEL CALCULATIONS

All experiments and calculations have been performed on vertical npn-transistors in common base configuration with a source impedance of 50 Ω .

The noise figures obtained refer to noise in the white spectrum, and noise figure is defined as the ratio between total noise power at the output of the device and the output noise power caused by the source resistor R_s [2] .

4a. Experimental results

Figure 5 shows a comparison between measurements and calculations of the transistor noise figure for a high frequency IC-transistor from process B (emitter dimensions 2.0 x 85 μm^2 ; α = 0.27).

For the calculations the value of the internal and external base resistances have been obtained from the lateral transistor dimensions and sheet resistances. The sheet resistances were measured on van der Pauw structures and the accurate lateral transistor dimensions were obtained by means of SEM photography. The DC and low frequency AC transconductance (G_m and g_m) and I_{0b} (indicated in fig.4) have been determined using simple DC transistor characteristics.

With these data we predicted the transistor noise figure as a function of collector current with the conventional noise model as stated in fig. 1, and with the modified model of fig. 4 where α = 0.27 has been substituted. It can be seen that the modified model accurately predicts the transistor noise behaviour (within 3%), whereas the conventional model predicts values that are too low in the medium and high current range.

4b. Numerical evaluation of transistor noise figure

Let us now evaluate the effect of sidewall injection on the noise performance for transistors with various emitter widths. As the effect of α on noise figure differs for different values of collector current density (see fig. 5), we defined a characteristic point in the noise figure curve to be used for our investigation.

We investigated the behaviour of the transistor noise figure as a function of emitter width for a constant collector current density J_c . For J_c we selected a value near the maximum F_t biasing condition; J_c = 200μA/μm²

Using geometrical scaling rules the relevant transistor model parameters were calculated for a number of different transistor geometries from process B. This resulted in a set of parameters for devices varying in emitter width from 0.4 μm to 5 μm The emitter length has been kept constant at 50 μm For the devices with emitter widths of 2.0 μm and more we used standard process B design rules. Model parameters for the transistors with narrower emitters have been obtained by applying design rules based upon a down-scaled version of the process.

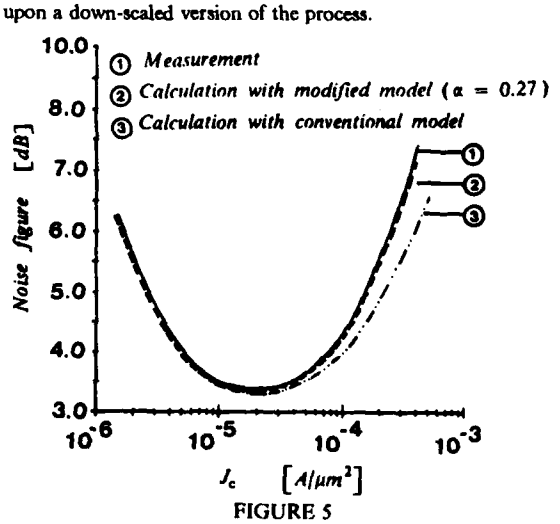


FIGURE 5

Transistor noise figure vs collector current density J_c for a transistor from process B (emitter dimensions 2.0x85 μm^2).

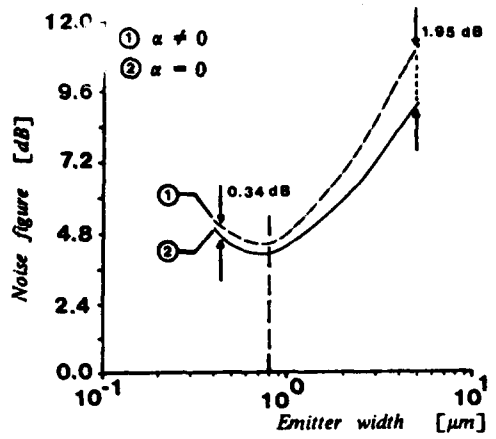


FIGURE 6

Calculated noise figure vs emitter width at collector current density $J_c = 200\mu A/\mu m^2$ for both the extended noise model ($\alpha \neq 0$) and the conventional model ($\alpha = 0$).

Figure 6 shows this characteristic noise figure calculated for $J_c = 200\mu A/\mu m^2$ as a function of emitter width. The calculations have been performed with the conventional model, where no sidewall injection has been taken into account, and with the extended model using values for α as plotted in fig. 3.

It can be observed that despite of the situation that wider emitters show a considerably lower value of α , the difference in calculated noise figure for wide emitters is much larger than for narrow emitters. This phenomenon can be explained as follows:

When the emitter width decreases, α increases, and we expect a strong effect of emitter sidewall injection on the transistor noise figure. However, when emitter width decreases, the internal base resistance decreases too. This means that the separation between bottom- and sidewall diode, as depicted in fig. 4, becomes less pronounced. As a consequence, regarding the noise performance, the influence of splitting up the e-b junction in a bottom- and sidewall component becomes less important for smaller emitter widths.

5. THE EFFECT OF SIDEWALL INJECTION ON BASE RESISTANCE CALCULATIONS

For a succesful determination of the total base resistance (R_b) from measurements of the white noise spectrum several arguments will have to be considered.

When the measurement is carried out at small collector current densities ($J_c < 1\mu A/\mu m^2$), the effect of collector shot noise tends to dominate the thermal noise of the base resistance and accurate extraction of R_b becomes difficult. A second problem in the low current region is the input impedance of the transistor. At small DC current levels the transistor input impedance may become very large with respect to the source resistance. This causes a large mismatch between the noise source and the transistor and may result in measurement errors.

For high collector current densities dissipation problems may occur, making it difficult to determine the actual transistor temperature. Also the same problem of mismatch becomes important.

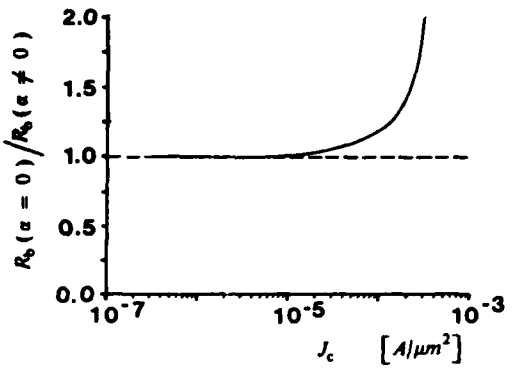


FIGURE 7

Ratio of R_b calculated back from noise figure using the conventional model ($\alpha = 0$) and R_b calculated back from noise using the extended model ($\alpha \neq 0$). The calculations have been performed on the transistor of process B with an emitter width of 2.0 μm

Applying the above mentioned arguments to our process B, we found that for a transistor with an emitter width of 2 μm and an emitter length between 10 and 100 μm , a practical current density region for the determination of R_b lies somewhere between 3 and 300 $\mu A/\mu m^2$.

However, when the influence of emitter sidewall injection on transistor noise figure is not taken into account, a large error in the calculation of R_b from measurements in this region may occur. This is illustrated in fig. 7, where we have plotted the ratio of R_b calculated back from noise figure using the model presented in fig. 1 (no sidewall injection incorporated) and by using the model as presented in fig. 4 (sidewall injection is taken into account).

When we conclude from fig. 5 that the results of the modified model (that takes into account the effect of sidewall injection) are quite accurate, we can interprete the curve of fig. 7 as the ratio between the R_b value obtained with the simple noise model (fig. 1) and the actual value of R_b .

6. CONCLUSION

A noise model for vertical bipolar transistors incorporating the effect of base current injection into the emitter sidewall has been derived. Comparison with measurements yields that for accurate noise figure calculations in the medium and high current range it is essential to take this sidewall injection into account.

REFERENCES

[1] G.A.M. Hurkx, IEEE Trans. El. Devices, to be published.
[2] A. van der Ziel, Solid State Physical Electronics, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1968.
[3] S.M. Sze, Physics of Semiconductor Devices, 1981.
[4] R.T. Unwinn and K.F. Knott, IEE Proc., vol. 127, Pt. 1, No. 2, April 1980.
[5] J.R. Hauser, IEEE Trans. Electron Devices, ED-11, pp. 238-242, 1964.
[6] G.Rey, Solid-State Electronics, vol. 12, pp. 645-659, Pergamon Press, 1969.
[7] J.A. Pals, Philips Res. Reports, vol. 26, pp. 91-102, 1972.
[8] H.K. Gummel and H.C. Poon, Bell Syst. Techn. J., vol.49, pp. 827-852, 1970.
[9] H.C. de Graaff, "Compact Bipolar Transistor Modelling," from Process and Device Modelling, W.L. Engl (editor), 1986.
[10] W.M. Webster, Proc. IRE, vol. 42, pp. 914-920, June, 1954.

CARRIER MULTIPLICATION AND AVALANCHE BREAKDOWN IN SELF-ALIGNED BIPOLAR TRANSISTORS

M.Reisch

SIEMENS AG, Central Research and Development,Microelectronics
Otto-Hahn-Ring 6, D-8000 Muenchen 83, FRG

1. Introduction

Consequent bipolar device miniaturization leads to an increase in epi layer doping and to a reduction in epi layer thickness. This implies an increase in maximum electric field strength in the BC junction for constant external voltages and has increased carrier multiplication in the BC diode as one of its consequences. As stable device operation demands the emitter collector breakdown voltage to exceed a critical voltage BU_{ceomin}, such carrier multiplication effects lead to important design considerations for scaling of bipolar transistors.

From the device characterization point of view measurement methods beyond the pure determination of breakdown voltages are needed. We have developped two very sensitive dc measurement methods which allow to obtain carrier multiplication data directly from the bipolar devices under study. We compare our measured results with calculations based on SIMS measurements and MEDUSA simulations, using GRANT's data [7], and find good agreement even for low values of applied base collector voltage if a 'dead space' correction is applied.

2. Measurement of electron multiplication factor Mn

Electron multiplication in reverse biased BC junctions is studied using a recently published [1] dc method (fig.1). A current source injects minority carriers through the forward biased EB junction into the base region. The carriers reach the reverse biased BC junction with a certain efficiency, denoted by α, and lead to an Ie induced current component that is proportional to the multiplication factor.

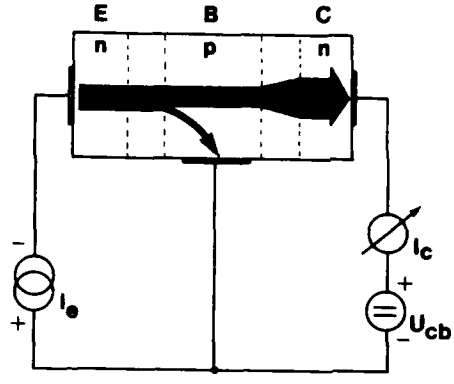


fig.1 Measurement setup for the determination of Mn

$$I_c(I_e,U_{cb}) = I_{cbo}(U_{cb}) + \alpha \cdot M_n(U_{cb}) \cdot I_e \qquad [2.1]$$

Mn is the value of the multiplication factor due to impact ionization of injected electrons. As shown in [1] differentiation of [2.1] with respect to the emitter current and division by α yields the multiplication factor Mn. The value of α may be determined by performing the differentiation procedure at zero applied base collector voltage - under this condition there is no multiplication in the BC diode.

Determination of α by the differentiation procedure is given preference over the extraction of α from the input and transfer characteristics of the device since great parts of the systematic measurement errors are compensated this way.

Under these circumstances the first order contributions of systematic measurement errors to the relative error in Mn are [9]

$$\frac{\Delta M_n}{M_n} = \frac{\delta}{\alpha \cdot \Delta I_e} \qquad [2.2]$$

where δ is only affected by the change in systematic error $\Delta Ic(Ic)$ for the different values of current Ie applied and vanishes as Mn approaches unity. Statistical contributions are not considered as principal limitations for the measurement accuracy since, by averaging, they die out as $1/\sqrt{N}$ with the number N of measurements performed. Therefore, for values of α close to unity high accuracy even in the range of low values of Mn(Ucb) may be obtained.

Carrier multiplication due to injected holes is characterized by the hole multiplication factor Mp(Ucb). In silicon Mp(Ucb) is well-known to be much lower than Mn(Ucb) if measured at the same voltage Ucb [2]. Application of our method for the determination of the hole multiplication factor is possible in principle, if holes are injected in the collector region using the substrate transistor, but as is shown by [2.2] the accuracy to be expected is poor due to the low value of the underlying α .

3. Photocurrent measurements

Virtually all experimental investigations [4..8] on carrier multiplication in silicon have been based on a chopped-light technique which in its basic form is due to CHYNOWETH and McKAY [3]. This and the possibility to obtain the multiplication factor Mp due to injected holes in addition to Mn are motivation to compare our results with results obtained from photocurrent measurements.

For our experiments, we investigate transistors without emitter metallization and measure the photocurrent $I_\omega(Ucb)$ in the base collector diode with the emitter potential left floating for different light intensities. $I_\omega(Ucb)$ is plotted vs. $I_\omega(Ucb=0)$, and the slope of the resulting straight line is determined by a least square fit (fig.2). Careful application of this method allows the determination of the 'optical' multiplication factor (identified with the slope of the straight line) for values larger than 1.001.

Due to the 'mixed injection' conditions, and the bias dependent extension of the BC space charge layer,

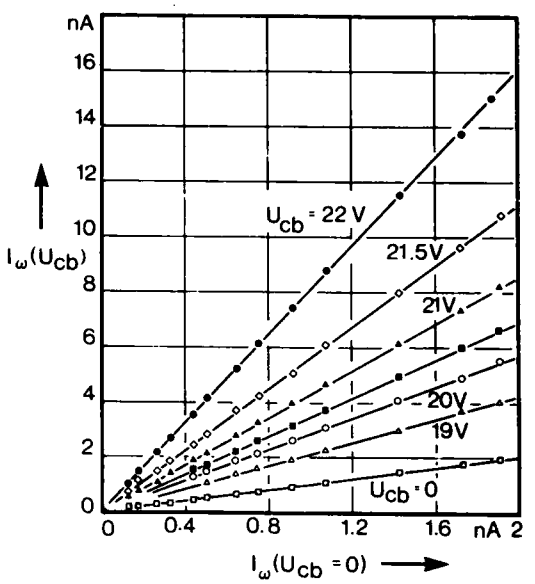


fig.2 Observed photocurrent $I_\omega(Ucb)$ vs. photocurrent $I_\omega(0)$ for Ucb=0

evaluation of Mp from these data is in general difficult. As long as a one-dimensional description applies, the hole multiplication factor may be obtained from

$$M_p(U_{cb}) = \frac{\partial I_\omega(U_{cb})}{\partial I_\omega(0)} \cdot f_p(U_{cb}) - M_n(U_{cb}) \cdot f_n(U_{cb}) \qquad [3.1]$$

Derivation of this formula is outlined in the appendix. fp(Ucb) and fn(Ucb) are weight functions depending on the 'minority carrier collection efficiencies' of the regions adjacent to the space charge layer as well as the bias dependent location of the space charge layer boundaries and the spatial distribution of the generation rate.

In fig. 3 we show the ratio of fn and fp vs. the absorption coefficient for an abrupt junction and a generation rate of the form

$$G_{opt} = \overline{G} \cdot \varkappa \cdot e^{-\varkappa x}$$

with the depth xjc of the metallurgical junction as parameter. This shows that calculation of fn, fp may be circumvented if the investigation is restricted to shallow EB structures and IR light sources ($\lambda \sim 830$ nm in our case). Under these conditions we may assume

$$\frac{f_n}{f_p} \ll \frac{M_n}{M_p}$$

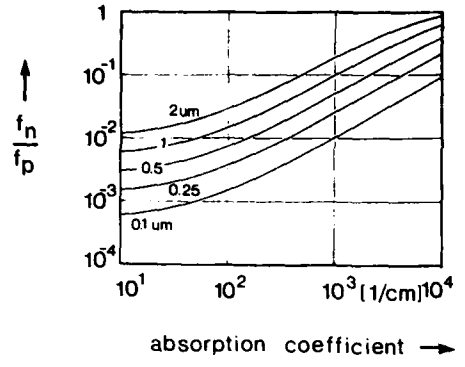


fig.3 Calculated ratio of weighting factors fn, fp for abrupt junction versus absorption coefficient with depth of metallurgical junction as parameter

(the hole current injected from the neutral collector region dominates) and identify the 'optical' multiplicatian factor with Mp.

Fig.4 shows measured values of Mn and $(\partial I_\omega(U_{cb}) / \partial I_\omega(0))$ vs. Ucb measured at the same device with a very shallow BC diode. Calculated ionization integrals for electron and hole injection derived from MEDUSA data, as discussed below are included for comparison.

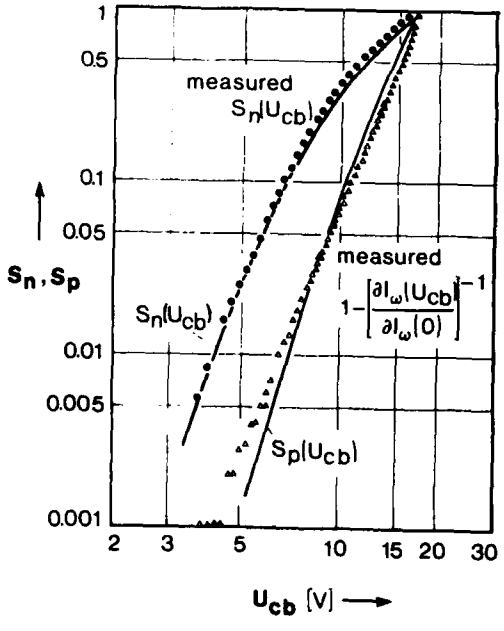


fig.4 Comparison of measured results for Sn = 1-1/Mn [●●●●●] and $1-1/(\partial I_\omega(Ucb)/\partial I_\omega(0))$ [▲▲▲,△△] with calculated Sn, Sp vs. Ucb behavior

4. Simulation of carrier multiplication behavior

We use doping profiles as determined by SIMS of measured transistors in MEDUSA simulations to determine the electric field distribution in the BC junction. Subsequent evaluation of the ionization integrals Sn(Ucb), Sp(Ucb) [10] allows a direct comparison of measured and calculated results.

Comparison of our measured curves to ionization integral calculations show that the ionization data of LEE et.al. [4], and GRANT [7] provide reasonable accuracy for BUcbo prediction. BUcbo values derived from van OVERSTRAETEN's [6] data appear to be too large by about 10 % for BUcbo ~ 20 V.

For voltages Ucb ~ BUceo the data of LEE provide the best description of the electron multiplication behavior observed if the ionization integral is evaluated without any corrections - but our measurements don't confirm the hole multiplication behavior predicted from LEE's data.

Considerable improvement in description may be achieved by introducing a simple 'dead space' correction [9]. This is performed by 'switching off' the electron and hole ionization coefficients in their respective 'dead spaces' defined by the

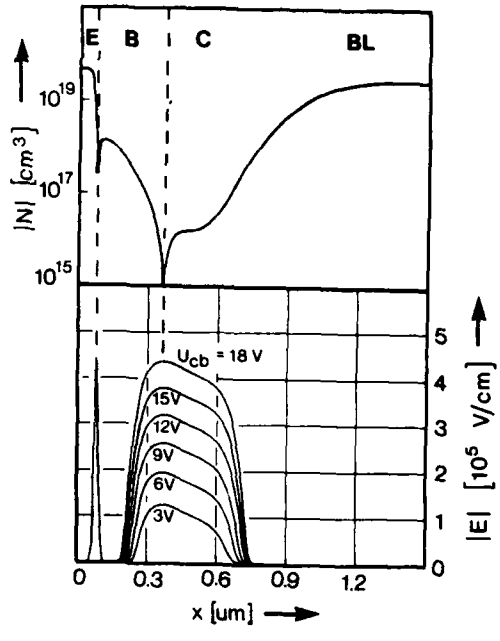


fig.5 Doping profile and electric field distribution in transistors with shallow BC diode

threshold energy of the corresponding impact ionization process.

In fig. 5 the vertical doping profile of the transistor of which the measured carrier multiplication behavior is plotted in fig. 4, is drawn together with the electric field distribution in the junctions, as obtained from MEDUSA, for different values of applied voltage Ucb.

These data for the electric field distribution and GRANT's data for the ionization coefficients were used, together with a threshold energy of 1.8 eV for electron induced impact ionization, to obtain the ionization integrals Sn := 1-1/Mn and Sp := 1-1/Mp [10] included in fig. 4. In spite of the crude approximation inherent in the assumption that the ionization coefficients are dependent on the local field strength only good agreement is obtained.

5. Conclusions

Our measurement method facilitates easy and accurate access to the electron multiplication factor Mn in the bipolar devices under test. Additional information is obtained from *dc photocurrent measurements*.

Carrier multiplication in the shallow BC diodes investigated may be described using ionization coefficients in CHYNOWETH form if a 'dead space' correction is performed.

Our methods provide the feedback necessary for the verification of simulated data and establish therefore a valuable tool for the optimization of the doping profile with respect to speed and breakdown.

Appendix: Photomultiplication

If there is no current injection from the boundaries the continuity equations [10] may be integrated to give

$$\frac{1}{q} j(x_c) = M_p(U_{cb}) \int_{x_b}^{x_c} dx\,[R-G] \exp(-\vartheta_n) \qquad [A.1]$$

with

$$\vartheta_n = \int_{x_c}^{x} d\bar{x}\,[\alpha_n - \alpha_p] \qquad [A.2]$$

and the hole multiplication factor Mp(Ucb).

Besides the photon induced current component the integral on the right hand side of [A.1] also takes account *of the 'dark current' which is due* to generation at SRH centers and interband tunneling.

We use [A.1] to derive for the slope of the I_ω(Ucb) vs. I_ω(Ucb=0) curve

$$\left.\frac{\partial I_\omega(U_{cb})}{\partial I_\omega(0)}\right|_{U_{cb}=\text{const.}} = M_p(U_{cb})\,\Theta_p(U_{cb}) \qquad [A.3]$$

where

$$\Theta_p(U_{cb}) = \frac{\left.\frac{\partial}{\partial G}\int_{x_b}^{x_c} d\bar{x}\,[R-G]\exp(-\vartheta_n)\right|_{U_{cb}}}{\left.\frac{\partial}{\partial G}\int_{x_b}^{x_c} d\bar{x}\,[R-G]\right|_{U_{cb}=0}} \qquad [A.4]$$

denotes a weighting factor unaffected by SRH generation and tunneling.

The space charge region divides the interval [xb,xc] in three subsections; the integrals in [A.3] may therefore be split up in three summands. We neglect recombination within the space charge layer, integrate the corresponding contribution by parts and use the mean value theorem [9] to obtain [3.1].

References:

[1] M.REISCH, IEDM 86 Technical digest, pp.654..657

[2] A.G.CHYNOWETH and K.G.McKAY, Phys.Rev. 108, 29 (1957)

[3] S.L.MILLER, Phys. Rev. 105, 1246 (1955)

[4] C.A.LEE, R.A. LOGAN, R.L.BATDORF, J.J.KLEIMACK and W.WIEGMANN, Phys.Rev.,134, pp.761..773 (1964)

[5] T. OGAWA, Jap.J.Appl.Phys. 4(7), 473 (1965)

[6] R.vanOVERSTRAETEN and H.deMAN, Sol.St. Electron. 13, 583 (1970)

[7] W.N.GRANT, Sol. St. Electron. 16, 1189 (1973)

[8] M.H.WOODS, W.C.JOHNSON and M.A.LAMPERT, Sol. St. Electron. 16, 381 (1973)

[9] M.REISCH, in preparation

[10] S.M.SZE, 'Physics of Semiconductor Devices',2nd Ed., John Wiley, N.Y. 1981

ELECTRICAL CHARACTERIZATION OF POLYSILICON/MONOSILICON INTERFACES

S. Bellone, P. Spirito,
University of Naples, Department of Electronic Enginnering
Via Claudio, 21, 80125 Naples

M. Arienzo
IBM T.J. Watson Research Center
Yorktown Heights, 10598 NY

The effective recombination velocity (ERV) associated with the polysilicon/monosilicon interface has been measured for different polysilicon structures. Using a new measurement method, the analysis indicates that the behavior of a polysilicon/monosilicon interface, free from any intentional oxide layer and with a polysilicon layer heavily doped, is similar to the one of a single-crystal high-low junction. It is demonstrated that blocking properties of a polysilicon contact improve if an undoped polysilicon layer is interposed between the doped polysilicon and the monosilicon when a significant arsenic concentration is present at the polysilicon/monosilicon interface.

1. Introduction

The interest of a polysilicon film in bipolar technology is related to the better blocking properties of the polysilicon/monosilicon interface, with respect to the one of the metal/single crystal interface, which can be thus used as emitter termination for higher current gain bipolar transistor. The mechanism responsible of the above improvement is generally attributed either to the tunneling phenomenon which can occur across a very thin oxide layer localized at the interface [1], or to a lower carrier mobility in the polysilicon layer [2] or finally, to a pile-up of the arsenic at the interface [3].

Usually the blocking properties of a polysilicon/monosilicon interface are evaluated by embedding the polysilicon layer in simple bipolar structures and extracting the current absorbed by this interface through its I-V characteristics. Since the current desired can result a negligible fraction of the total current, the contribute of the other regions of the test structure must be evaluated separately.

By using a new measurement technique [4] based on a different test structure, which does not require the knowledge of any critical parameter, we report on experimental results obtained on different polysilicon/monosilicon interfaces. A comparison with a metal contacted implant layer shows that the behavior of the polysilicon/monosilicon interfaces investigated can be described by means a high-low junction model.

With the help of the results in [5] our findings confirm that the blocking properties of a polysilicon/monosilicon interface strongly degrade at the highest doping level of the polysilicon layer. However, since heavy doped polysilicon layers are required in order to reduce the series resistance value, it is demonstrated that the blocking properties of polysilicon/monosilicon interfaces can be improved by interposing a thin undoped polysilicon layer between the monosilicon and the heavily doped polysilicon layer.

2. Experimental

The measurement method requires the test structure sketched in fig. 1. As described in [4] it consists basically of a P^+-N-N^+ injecting diode between surface and the substrate, with an added N^+ region, in this case the polysilicon N^+ layer under test, surrounded by the P^+ region at a distance much smaller than the diffusion length in the epilayer. The polysilicon layers were deposited on two different silicon epitaxial wafers 1μm thick and arsenic doped $2x10^{15}$ cm^{-3} or $1x10^{16}$ cm^{-3} respectively, using a CVD reactor at 670 °C. Before polysilicon deposition, all the wafers were done a dip etch in buffered hydrofluoric

The work was supported by Italian Research Council under the program:
"Materiali e Dispositivi per l'Elettronica a Stato Solido"

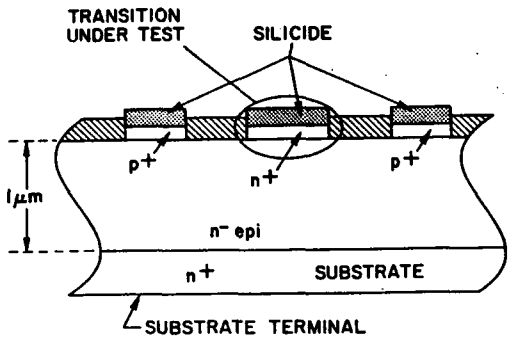


Fig.1 Schematic of the test structure

acid (BHF) to remove the oxide layer on the surface [6]. After the poly deposition, all the wafers received a total heat cycle of 820 °C for 160 min during the subsequent thermal processes required by the realization process. All the structures were contacted by using Titanium silicide formed at 700 °C and by removing the unreacted titanium by etching.
In particular, three different polysilicon layers were fabricated on different wafers:

a) 150nm in-situ arsenic doped, further annealed at 900 °C/ 15min
b) 150nm in-situ arsenic doped
c) 150nm of undoped film followed by 150nm of in-situ arsenic doped (bilayer)

To make a comparison with a single crystal layer, the same test devices were fabricated by implanting the N^+ and P^+ surface regions directly on the silicon layer using respectively phosphorous, with energy 50KeV and dose $1x10^{15}$ cm^{-2}, and boron, with energy 20KeV and dose $1x10^{15}$ cm^{-2}, on the same epitaxial layer as above. Such devices were annealed at 900 °C/ 20min and then contacted with Al-Cu alloy. In fig.2 the doping profile obtained using a CAMECA Secondary Ion Mass Spectroscopy (SIMS) system is presented for both the polysilicon layer type (a) and the N^+ implanted layers.

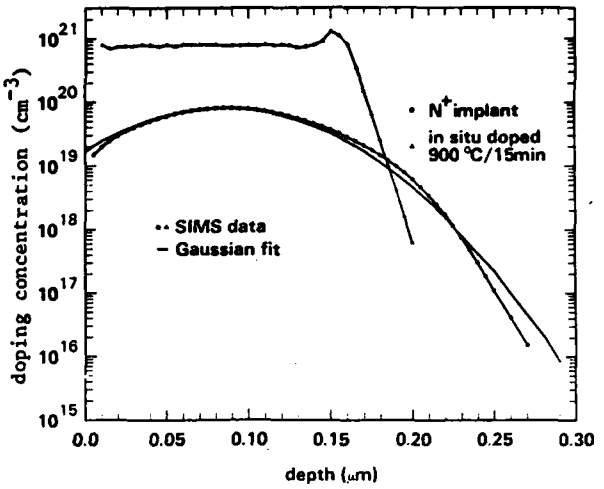


Fig.2 Doping profiles obtained from SIMS meas.

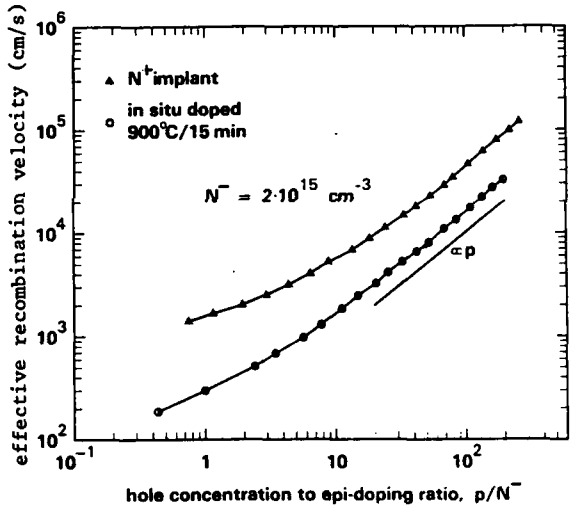


Fig.3 Experimental dependence of the ERV as a function of the injection of minority carrier

3. Results and discussion

A typical experimental result of the effective recombination velocity for the polysilicon structure type (a) and the single crystal implant layer are reported in fig. 3 as a function of the minority carrier injected in the epilayer. As it can be observed, both ERV curves vary with the minority carrier density according to following equation [4]:

$$S \approx S_r + S_n\,(1 + p/N^-) \qquad (1)$$

allowing the extraction of the S_r and S_n components respectively from the low and the high injection branch of the curve. In the above equation N^- is the epi doping, while S_r and S_n describe respectively the recombination component of the space charge region and the heavily doped region of the transition. The latter one for a metal contacted implant layer is given by [7]:

$$S_n = \frac{N^-}{\int \frac{N_{Deff}(x)}{D_p}\,dx} \qquad (2)$$

where the integral term is extended over the layer portion with a doping greater than $1x10^{18}$ cm^{-3} [8].

The mean values of both S_n and S_r components measured at least on thirty devices for each epilayer doping, are summarized in Table I for the polysilicon structures type (a) and the implant layers. The above values of S_n for the implant layer are in good agreement with the theoretical evaluation from Eq. (2) using for the doping profile the best gaussian fit of the SIMS data in fig.2, the minority carrier mobility as in [9], and the bandgap narrowing according to Slotboom and De Graaff model [10].

It is worth noting from Table I that the numerical value of ERV of the polysilicon layer in the low injection limit depends on the contribution of both components S_n and S_r, with the former one prevailing as the epi doping increases, while in the case of the implanted region the component S_r dominates always. Such difference can be explained by means of the greater sharpness of the doping profile for the polysilicon layer, as shown by the SIMS data in fig.2, and of the demage introduced by the implant process. It can be useful to describe the blocking properties of the poly by means the leakage current [7]:

$$J_{so} = q \frac{S_n\, n^2_{io}}{N_{Depi}} \qquad (3)$$

J_{so} in our case results $3x10^{-12}$ Acm^{-2}. This value is in agreement with the trends reported by [5] for the dependence of J_{so} on poly doping and confirms that the blocking properties of a polysilicon/monosilicon interface strongly degrade as the doping level of the polysilicon layer becomes greater than $2x10^{20}$ cm^{-3}.

TABLE I

Experimental values of S_n and S_r for the implanted and polysilicon layers, at different epi dopings N^-. The standard deviations are reported into parenthesis.

Epi doping		N^+ Implant	In-situ doped 900 C/15 min
$2x10^{15}$	S_n	434 (14)	161 (12)
	S_r	1321 (393)	151 (48)
10^{16}	S_n	2445 (150)	667 (43)
	S_r	2380 (384)	< 200

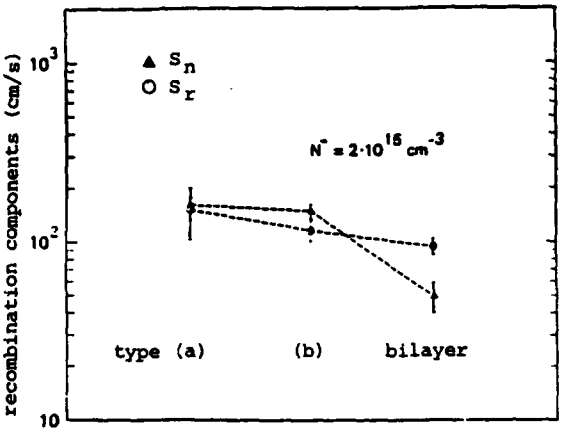


Fig. 4. Experimental results of the ERV components for all the polysilicon structures

Since heavy doped polysilicon layers are normally required in order to reduce the series resistance value, we examined the effect of a lower arsenic concentration at the polysilicon/monosilicon interface by realizing the bilayer structure, where an undoped polysilicon layer is interposed between the monocristalline and the in-situ doped layer.

The results are summarized in fig. 4 , where a significant lowering of the total ERV can be observed in the bilayer structure, as it has been already noted in [11]. To be sure that above behavior is not due to the different thermal cycle of the bilayer structure, the same in-situ doped layers were deposited directly on the monocristalline silicon to realize the structure type (b) and then processed identically as the bilayer. As shown in fig.4, the S_n component of the in-situ doped polysilicon layers does not change significantly with the heat treatment, indicating that the doping activation was the same for the two different time-temperature values of annealing.

In order to understand better the above behavior of the ERV components, the temperature dependence of S_n for all devices is reported in fig. 5. The theoretical curve drawn for the

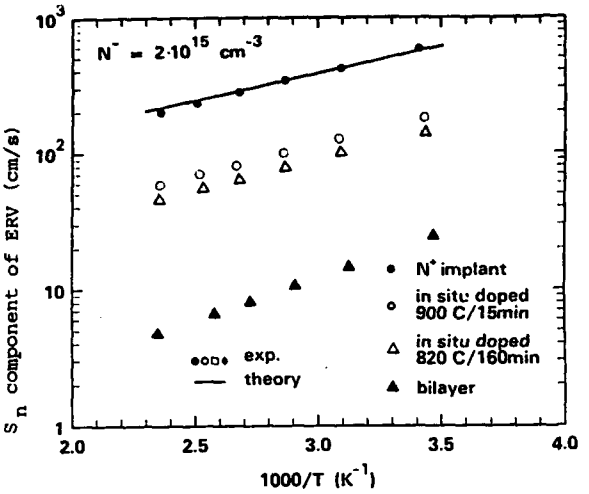


Fig.5 Temperature dependence of the S_n components for all the structures investigated.

case of the implanted region has been obtained using Eq. (2) and assuming the minority carrier mobility independent of the temperature. The similar temperature dependence of the polysilicon layers type (a) and (b) confirms that the two thermal treatments have no substantial effect in altering the physical parameters and hence the recombination velocity of the polysilicon/monosilicon interface, as it has been already noted in fig. 4. Moreover the strict correlation between the results for the implant and the polysilicon layers in Table I and fig. 5 is a further proof that the behavior of our polysilicon structures can be described through high-low junction model. If we attribute the overall temperature dependence to the bandgap narrowing phenomenon described as in [10] , the slope for the above mentioned polysilicon structures supports an activated doping value $N^+ = 6 \times 10^{19}$, which agrees with the results of [12].

Surprisingly, the temperature dependence of the bilayer structure in fig. 5 shows that the reduced recombination S_n is associated with a greater temperature dependence. A possible explanation of the bilayer behavior can be given recalling that during the thermal cycle arsenic diffuses in the undoped polysilicon layer through the grain boundaries, while the polysilicon grains remain less doped. As a result of this preferential grain-boundary diffusion, a thin heavily doped interfacial region is formed at the polysilicon/monosilicon interface, giving rise to an additional barrier for the minority carrier injected into the polysilicon layer. Then the undoped layer is characterized by a greater lifetime into the grain, due to the lower doping level, and by a better blocking high-low junction, created between the undoped grain and the grain-boundary, where the atoms segregation is enough to reduce the amount of defect centers [5].

4. Conclusions

Comparing the blocking properties of the polysilicon/monosilicon interface with that of a single crystal high-low junction, we have shown that the high-low junction model is still applicable to the polysilicon junction in the case of a high polysilicon doping. The presence of a thin undoped polysilicon layer between the silicon and the in-situ doped polysilicon seems to further improve the blocking properties of a polycristalline contact.

References

[1] H.C. De Graff, and J.G. De Groot, IEEE Trans. Electron Devices, ED-26, (1979), 1771

[2] T.H. Ning, and R.D. Isaac, IEEE Trans. Electron Devices, vol. ED-27,(1980) 2051

[3] H. Ryssel, H. Ilberl, L.Bleier, G, Prinke K. Haberger, Appl. Phys., vol. 24, (1981), 197

[4] S. Bellone, A. Caruso, and G.F. Vitale, IEEE Trans. Electron Devices, (1985) 1771

[5] G.L. Patton,J.C.Bravman, and J.D. Plummer, IEEE Trans. Electron Devices, (1986) 1754

[6] P. Ashburn and B. Soerowirdjo, IEEE Trans. Electron Devices, ED-31, (1984) 853

[7] M.A. Shibib, F.A. Lindholm, and F. Therez, IEEE Trans. Electron Devices, (1979) 959

[8] J.G. Fossum, and M.A. Shibib, IEEE Trans. Electron Devices, vol. ED-28, (1981) 1018

[9] S.E. Swirhun, J.A. Del Alamo, and R.M. Swanson, IEEE Electr. Dev. Lett., (1986) 168

[10] J.W. Slotboom and H.C. De Graaff, Solid-State Electron., vol. 19, (1976) 857

[11] A. Neugroscel,M. Arienzo,Y. Komen,and R. Isaac,IEEE Trans. Electron Devices,(1985) 807

[12] M. Arienzo, A.C. Megdanis,P.E. Sackles, IEEE Trans. Electr. Dev.,(1986) 1535

CURRENT GAIN DEPENDENCE ON THE EMITTER SIZE OF POLYSILICON-EMITTER BIPOLAR TRANSISTOR

M. Miura-Mattausch

Siemens AG, Corporation Research and Development, Munich, FRG

It is shown that the current gain increases for reduced emitter sizes of self-aligned bipolar transistors with polysilicon emitter contact. This phenomenon is explained by the difference between the polysilicon contact area and the effective emitter area derived from electrical data.

1. INTRODUCTION

In order to achieve high-speed ICs, technologies are evolving in the direction of reduced critical dimensions and reduced parasitic elements. Present trends in the bipolar technology emphasize polysilicon layers both as diffusion sources and to provide self-aligned contacts[1]. Beside the reduction of the inactive base region and of the emitter depth, this new technology exhibits new phenomena. The current gain ß can be much higher than for the conventional case[2]. A significant tunneling is sometimes observed in the base current I_B[3].These phenomena have been explained either by the low carrier mobility in the polysilicon or by the existence of an interface. In addition to these phenomena, an unexpected dependence of ß on the emitter size is shown; ß increases with reducing the emitter size. To understand this phenomenon, ß is measured for several differently prepared npn transistors. The transistors with different emitter depth are obtained by only varing the emitter diffusion temperature, keeping the polysilicon thickness and the polysilicon doping concentration the same. With the help of a theoretical concept, it is shown that the phenomenon is explained by the special features of the polysilicon-emitter contact. For an economic circuit design ß should not be dependent on the emitter size. The consequences for the design of a polysilicon technology are discussed.

2. MEASUREMENTS OF THE CURRENT GAIN ß

Figs1a and 1b show ß vs. V_{EB} for different emitter sizes with two emitter-diffusion temperatures; 950° and 900°C. SIMS measurements give an emitter depth X_E of about 150nm and 80nm, respectively. Though ß is usually shown as a function of the collector current I_C, it is shown here as a function of the emitter-base bias V_{EB}. With this representation the dependence of I_B on ß as well as the bias dependence can be seen. The following features are recognized for the ideal region of I_B,I_C:

(a) ß increases with reducing the emitter size,
(b) the tendency is enhanced for the 900°C case,
(c) in the 900°C case ß(max) is shifted to higher V_{EB} with smaller emitter size.

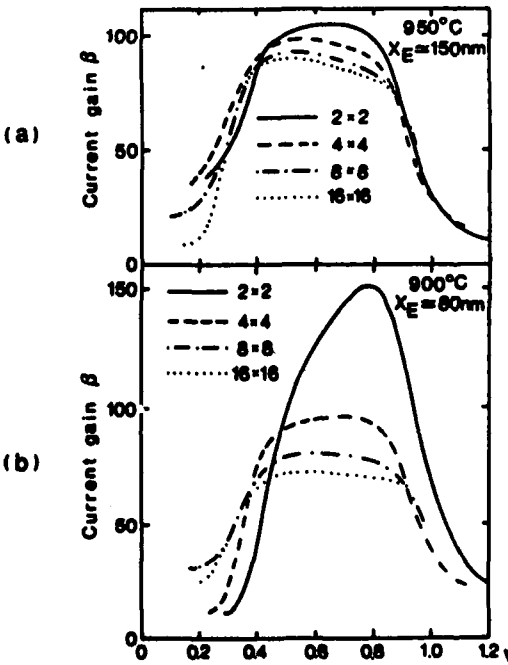


Fig. 1. ß vs. V_{EB} for two emitter-diffusion temperatures. Emitter mask sizes are in μm^2.

Features (a) and (c) are recognized also for small stripe transistors ($2x4/\mu m^2$, $2x8/\mu m^2$,etc.). The emitter size is described by the mask size, which is larger than the real emitter contact. According to the reduction of the emitter size, a reduction of ß is expected because of the large contribution of the peripheral current to I_B (see Fig.2)[4].Our measurements show the opposite tendency; the increase of ß.

From a comparison of two $2x2/\mu m^2$ transistors with different diffusion temperatures, it is concluded that the increase of ß is mainly due to the reduction of I_B, which may be attributed to the effect of the interface. If the emitter is deep, injected holes from the base to the emitter recombine with electrons in the bulk and are scarcely influenced by the interface. The TEM pictures show that the interface oxide layer is a little bit thinner than 1nm for the 900°C case, and is broken and the recrystallization starts for the 950°C case.

3. THEORETICAL ANALYSIS

3.1. Characterization of polysilicon contact

Due to the reduction of x_E, holes do not completely recombine in the bulk, but arrive at the interface. The hole current density at the monosilicon side of the interface is modeled by the surface-recombination velocity S_p[5]

$$j_p(x=0) = qS_p\Delta p(x=0), \quad (1)$$

where q is the electron charge and $\Delta p(x=0)$ is the excess hole density at the interface. The coordinate x denotes the direction vertical to the emitter surface. The effects of the polysilicon layer and the tunneling may be lumped together in S_p, where the recombination velocity at the interface is the dominant factor at bias voltages up to intermediate values.

3.2. 1D approximation

A recombination velocity in the emitter is introduced in the same way as S_p

$$j_p(x) = qS(x)\Delta p(x). \quad (2)$$

Using the continuity equation for holes with the boundary condition $S(0)=S_p$, the recombination velocity at the emitter-base junction $S(x_E)$ can be solved. With approximations that the doping profile is a step function, and that the diffusion coefficient D_p as well as the diffusion length L_p are constant within the emitter,

$$S(x_E) = v_d\frac{v_d+S_p\coth(x_E/L_p)}{v_d\coth(x_E/L_p)+S_p}, \quad (3)$$

where $v_d=D_p/L_p$. Eq.3 shows that the injected minority carriers into the emitter can be described by a single additional parameter $S(x_E)$ to the conventional model parameters

$$j_p = qS(x_E)\frac{n_i^2}{N_{Deff}(x_E)}\exp(\frac{qV_{EB}}{kT}), \quad (4)$$

where n_i is the intrinsic concentration and $N_{Deff}(x_E)$ is the effective doping density including the bandgap narrowing.

3.3. 2D contribution to the current

By reducing the emitter size the ratio of the peripheral contribution to the emitter area increases. As a result a 2D peripheral current may play an important role (see Fig.2). An analytical solution for estimating the current within a 2D treatment has been shown by reducing the continuity equation to a pair of dual integral equations with cylindrical coordinates[6].We

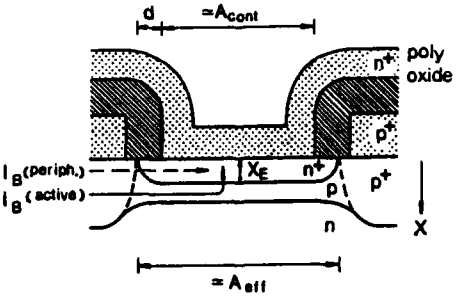


Fig. 2. The schematical cross-section of the transistor.

apply the treatment to the hole flow in the emitter by assuming the emitter-base junction to have a circular shape. The result is shown in Fig.3. It shows the current ratio of the 2D to the 1D treatment vs. the device radius r divided by L_p for several different ratios of S_p to v_d. For smaller S_p the 2D contribution becomes more enhanced. Usually the polysilicon contact with

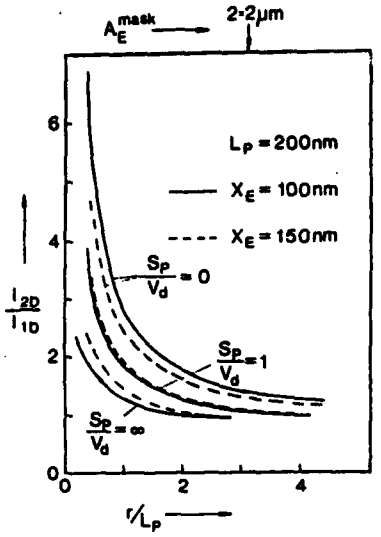


Fig. 3. The current ratio of the 2D to the 1D treatment vs. the device radius r with different S_p values.

1nm oxide layer at the interface exhibits S_p/v_d=1 In this case the 2D contribution should be imptant for structures smaller than 2x2/µm^2. However for the interface with oxide layer thicker than 1nm ($S_p \approx 0$) the 2D contribution is not negligible already for the 2x2/µm^2 emitter size. The dependence of 2D contribution on x_E is depicted also in Fig.3. For emitters larger than 2x2/µm^2 and interface oxides below 1nm thick, the 1D approximation may be sufficient.

3.4. ß dependence on the emitter size

Fig.4 shows the estimated S_p vs. the effective emitter area A_{eff} derived from electrical data at V_{EB}=0.6V under the 1D approximation for the case of Fig.1a. Since I_C is less influenced by the interface condition and has much less peripheral contribution than I_B, A_{eff} is estimated from I_C . As seen in Fig.2, A_{eff} is not identical with the geometrical emitter contact area. As parameters standard values for device simulation are used[7].S_p is almost a linear function of A_{eff}. It is hard to recognize any change of the interface condition according to the different emitter sizes in TEM pictures. We consider that the linear relation may come from

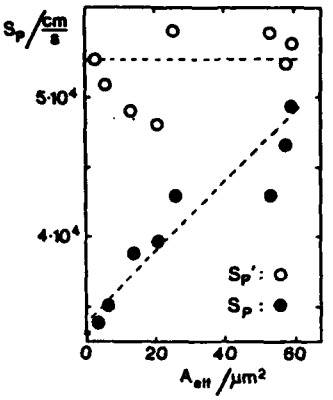


Fig. 4. S_p vs. A_{eff}.

the difference between the emitter contact area A_{cont} and A_{eff}. A_{cont} is expected to be smaller than A_{eff}, because of the lateral diffusion under the oxide spacer. Since a homogeneous current density is available in the whole A_{eff} under the 1D approximation, the average S_p is simply written

$$S_p = \frac{S_p' A_{eff} + S_p'' A_{insu}}{A_{eff}}, \qquad (5)$$

$$A_{insu} = A_{eff} - A_{cont},$$

where S_p' and S_p'' are the surface-recombination velocities for the polysilicon contact and for the spacer areas, respectively. The estimated S_p' by assuming S_p''=0 is also depicted in Fig.4. S_p' is not size-dependent but is constant as we expect. The length d (see Fig.2), which describes approximately the lateral diffusion underneath the spacer, can be determined from the difference between A_{cont} and A_{eff}. The result is d 170nm for the x_E=150nm case.

The fact that the geometry dependence of ß is enhanced for the shallow emitter can be explained within the same theory. The dependence of $S(x_E)$ on x_E is examined at V_{EB}=0.6V. The result is shown in Fig.5. By reducing x_E an enhancement of the size effect can be seen from the steep gradient of $S(x_E)$ vs. A_{eff}.

As seen in Fig.1b the steep increase of ß as a function of V_{EB} has two components: a first increase and a second increase. The geometry

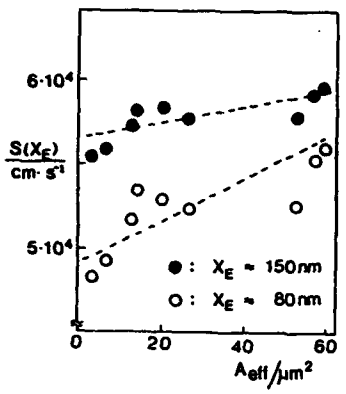


Fig. 5. The recombination velocity at the emitter-base junction $S(x_E)$ vs. A_{eff}.

dependence of the first increase has been explained by the difference between A_{cont} and A_{eff}. The extreme increase of ß for small and shallow structures is owing to the second increase, which also leads to a shift of the maximum of ß to higher value of V_{EB} and to the disappearing of the favorable plateau. In this case a homogeneously distributed thin oxide layer, about 1nm thick, at the interface is recognized by TEM pictures. The second increase of ß is due to a further reduction of the injected current in I_B. To confirm whether the reduction of I_B in the second increase is due to the exsistence of the interface, the series resistance is estimated from the drop of I_B in this region, where the applied voltage is still below the high injection region. The result is shown in Fig.6. At low current densities the resistance is not constant but depends on the current for small

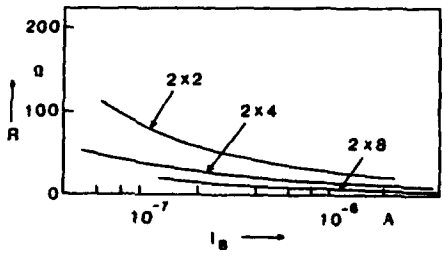


Fig. 6. The estimated emitter series resistance R vs. A_{eff} for different emitter mask sizes given in µm².

sizes. The dependence disappears rapidly as increasing the emitter size. As the size of the emitter mask area A_m is reduced, the emitter series resistance increases approximately as a function of $1/A_m$. This relationship suggests that the interfacial resistance dominates in the estimated series resistance.

5. CONCLUSION

It has been shown that ß exhibits a geometry dependence for self-aligned transistors with the polysilicon emitter contact, i.e., ß increases with reducing the emitter size. This is explained by the difference between the polysilicon contact area and the effective emitter area derived from electrical data. For a convenient circuit design it is important to reduce the lateral diffusion of the emitter underneath the oxide spacer. The most serious problem of reducing the emitter size and the emitter depth is the interface resistance. To decrease the contact resistance, the reduction of the thickness in the interfacial oxide layer is important.

ACKNOWLEDGEMENTS

The author acknowledges helpful and valuable discussion with L. Treitinger, H. Schaber, T. Meister, R. Schreiter, and H. Caver.

REFERENCES

[1] T.Sakai,Y.Kobayashi,H.Yamaguchi,M.Sato,T. Makino,Jap.J.Appl.Phys.(Suppl.),20,155(1981).

[2] T.H.Ning and R.D.Issac,IEEE Trans.Electron Devices,ED27,2051(1980).

[3] H.C.deGraaff and G.deGroot,IEEE Trans. Electron Devices,ED26,1771(1979).

[4] H.-M.Rein,Solid-State Electron.,27,625(1984).

[5] Z.Yu,B.Ricco,and R.W.Dutton,IEEE Trans. Electron Devices,ED31,773(1984).

[6] P.-T.Chen,K.Misiakos,A.Neugroschel,and F.A.Lindholm,IEEE Trans. Electron Devices,ED32,2292(1985).

[7] B.Benna,T.F.Meister,and H.Schaber, to be published in Solid-State Electron.

Session C4.1

Test Chips

Chairman: J. Robertson

Thursday, September 17, 1987

C4.1.1

The effect of device geometry on IGFET characteristics.

J. A. Serack, A. J. Walton, J. M. Robertson

Edinburgh Microfabrication Facility, University of Edinburgh, Edinburgh, Scotland.

A novel technique for the fabrication of asymmetrical incompletely gated transistors is described. The subthreshold characteristics of transistors fabricated using the technique are measured and conclusions concerning the mechanisms responsible for the observations are presented.

1. INTRODUCTION

Process designers have been using the primary relationship between gate capacitance and gate overlap of the source and drain for decades to maximise circuit speed by reducing gate capacitance. However little attention has been paid to how the correspondly reduced gate overlaps effect the D.C electrical characteristics of the transistors.

Recently asymmetries in transistor characteristics have been encountered [1-3] when the source and drain connections of the transistor were reversed. Gaps in the gate to channel coverage at either the source or drain end of the channel have been identified as the cause of the asymmetries. The gaps arise from the gate shadow cast during off-axis implantation, (which is universally employed to reduce channeling), coupled with a low thermal budget for the subsequent steps of the process. In order to study the effect that these gaps have on transistor characteristics, sidewall spacers have been used during source-drain implants together with variations in the drive-in times, to obtain transistors with variations in gap size [2].

Unfortunately other drive-in time dependent factors, such as the source and drain depth or the degree of channel implant activation, cause variations in the electrical characteristics of transistors fabricated with the above technique. Unpredictable variations between wafers also makes a controlled comparison between transistors with different magnitudes of gaps or overlaps impossible. The uncontrollable factors inherent in this approach make it only useful for obtaining a qualitative understanding of the effect. This paper presents a novel technique for the fabrication of transistors with a predictable range of gaps and overlaps upon a single wafer. Measurements performed on transistors constructed using this method are free from the factors discussed above.

2. EXPERIMENTAL SET-UP

In order to control the degree of gate overlap in this experiment the convenience of self-aligned gates had to be forgone and separate steps used to define the gate and channel. A fairly typical LOCOS isolation process was employed until after the source and drain implants were completed. What would have then been the polysilicon gate was stripped away and the contact holes cut. Aluminium was patterned by reactive ion etching to form the gate and contacts to the source and drain. The process used <100> silicon wafers and resulted in the following features; arsenic source-drains 0.35 μm deep, a channel impurity concentration of 8×10^{16} atoms/cc of boron, 600 Å thermal gate oxide, and 0.5 μm thick aluminium gates.

Although excellent alignment errors of less than 0.3 μm for the gate to channel were obtained using a 10:1 reduction direct to wafer stepper for the photolithography, the small gate gaps and overlaps are dictated by the chip design.

The chip layout, shown schematically in figure 1, was realised using a relational database to first define a single transistor and then replicate it to form an array of transistors. Then algorithms were applied to adjust both the channel and gate sizes and the gate to channel alignment at each site. This resulted in an array of transistors whose nominal channel lengths vary across a row from 1.15 to 2.05 μm and whose gate to channel alignment varied down a column from 0.75 μm drain gaps to 0.45 μm source gaps, in steps of 0.15 μm. Reference transistors with 5 μm drawn channel lengths and 0.3 μm overlaps were also included. The choice of 0.15 μm as the step size was made after a previous experiment showed the metal edge roughness was of this magnitude for 0.5 μm thick aluminium.

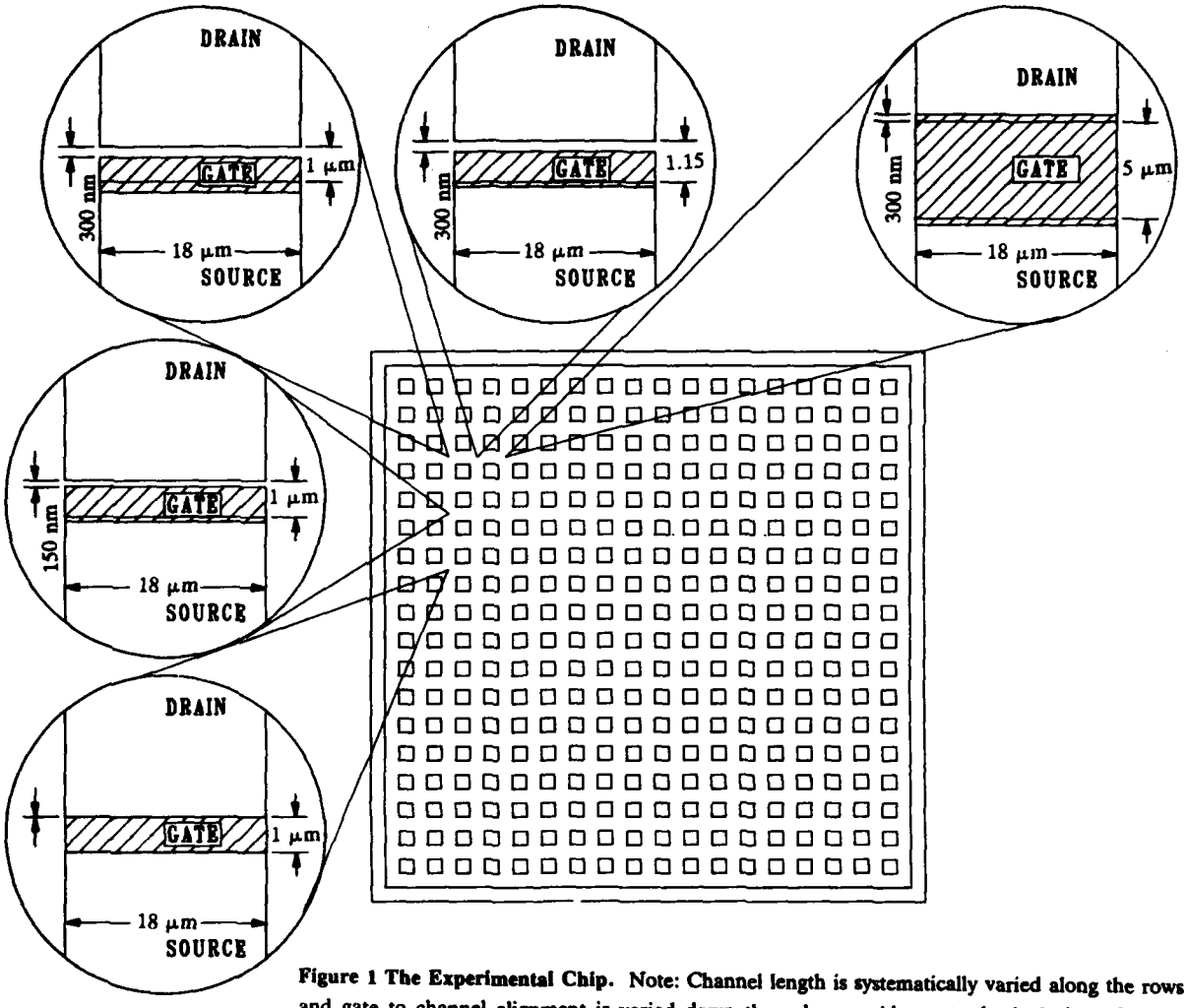


Figure 1 The Experimental Chip. Note: Channel length is systematically varied along the rows and gate to channel alignment is varied down the columns. Also note the inclusion of 5 μm reference transistors.

3. SITE SELECTION and MEASUREMENT CONSIDERATIONS

Once a particular channel length column on a die was selected for characterisation, alignment identification was accomplished by comparing both the normal and source-drain reversed subthreshold curves of each transistor in the column. The transistor or transistors that have coincident curves in both configurations have complete gate coverage of the channel. Figures 2 and 3 show this as well as other characteristics for two transistors from a column. The transistor in figure 2 is incompletely gated with a large drain gap which is evident in the asymmetry of the subthreshold curves (figure 2a). The transistor in figure 3 is from a couple of sites farther down the same column and is aligned. Comparisons of the other characteristics show similar results to that reported in reference [2].

Once the subthreshold curves have been used to determine the first aligned transistor in the column, the structure of the layout can be used to select transistors with a particular gap or overlap for further measurement.

One important consideration when measuring transistors with gaps is their sensitivity to charge accumulation over the gap in gate to channel coverage. If the gap is uncovered water vapour can be charged by the gate and the transistor can appear to change into a non-gapped transistor after a few measurements. Even with a passivation layer the gaps sensitivity to charge requires that any charge added by the gate is removed. This can be accomplished automatically by the measurement software reversing the sense of the gate voltage approximately every quarter second. Of course observations are only recorded during the positive sense.

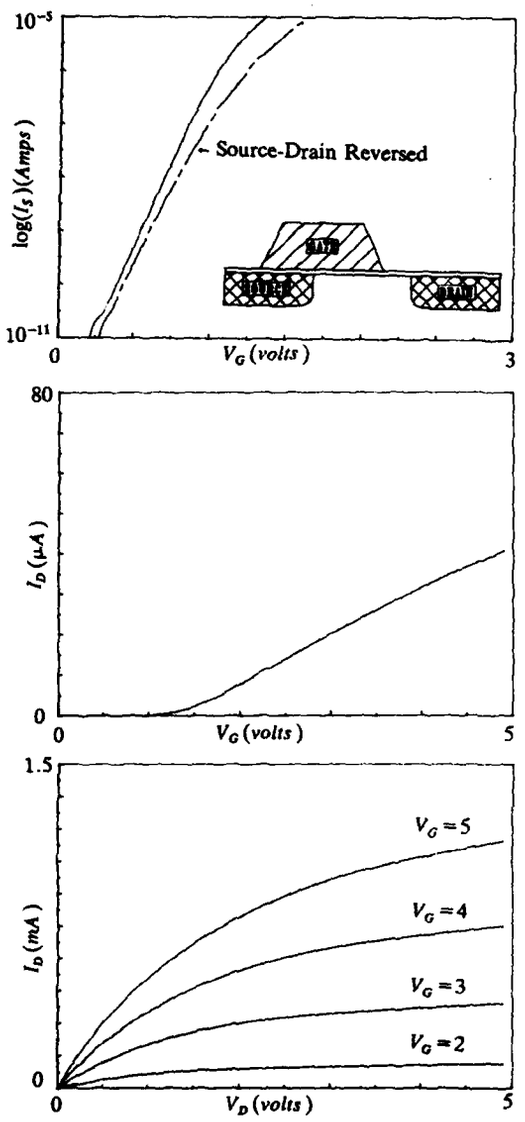


Figure 2 Electrical Characteristics of a 1 μm long transistor with a 0.2 μm drain gap. Figures a-c (top to bottom) are in the subthreshold, threshold and saturation regions.

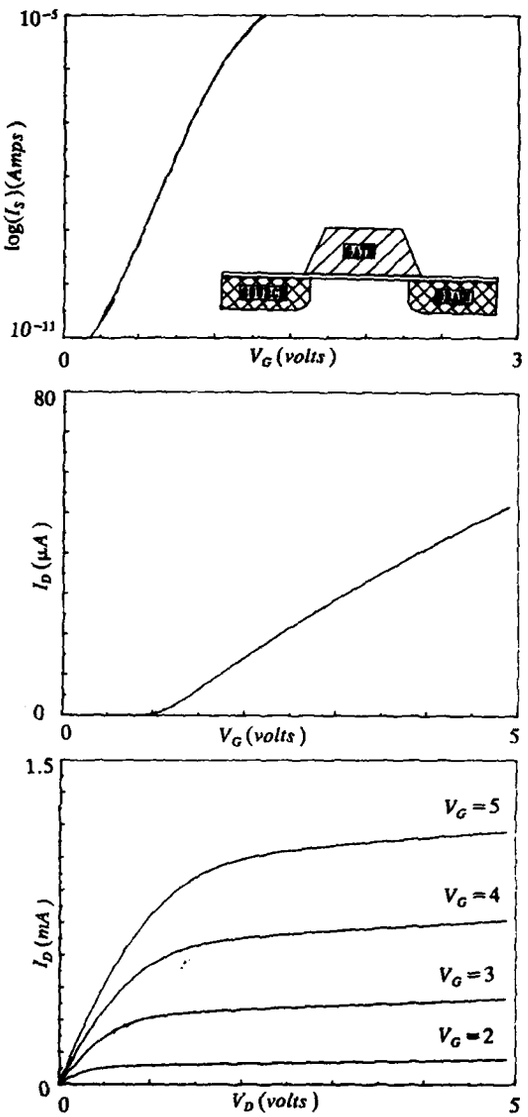


Figure 3 Electrical Characteristics of a 1 μm long transistor with complete channel coverage. Figures a-c (top to bottom) are in the subthreshold, threshold and saturation regions.

4. OBSERVATIONS - SUBTHRESHOLD REGION

An important parameter in the sub-threshold region of operation is the subthreshold swing [4]. It is a measure of how large a change in gate voltage is required to change the channel current by one order of magnitude. Figure 4 is a plot of the sub-threshold swings dependence on the drain voltage for each transistor in an alignment column.

Two generalisations are immediately evident.

Those transistors that have a gap in gate to channel coverage near the source end of the channel have an increased swing that is independent of drain voltage. If the swing for these source gapped transistors (SGT's) is plotted relative to the magnitude of the gap a straight line results. That relationship may be useful in analysis of the gate fringe fields.

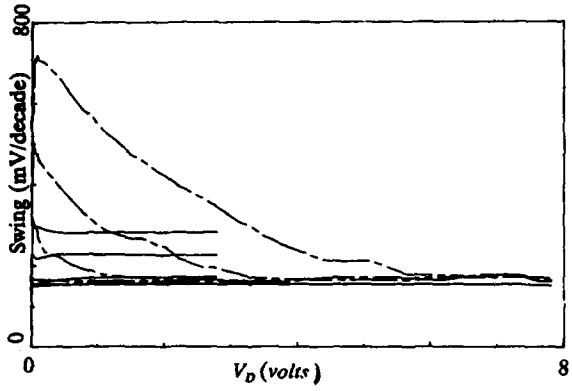


Figure 4 Subthreshold swing as a function of transistor structure and drain voltage. The dashed curves are for DGT's with gaps of (top to bottom) 0.50, 0.35, 0.20, and 0.05 μm. The solid curves are for SGT's with gaps of (top to bottom) 0.35, 0.20, and 0.05 μm. Note: The DGT's are drain voltage dependent.

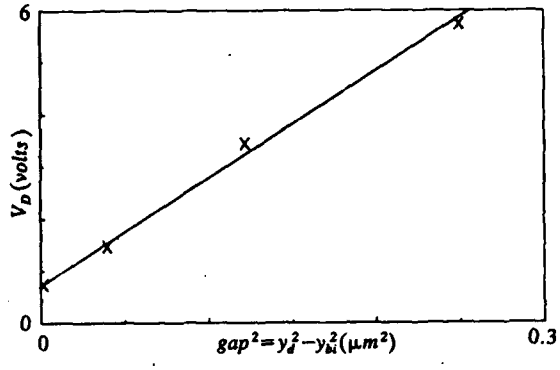


Figure 5 The drain voltage causing constant subthreshold swing as a function of the magnitude of the DGT's gap. Note: The linear dependence supports equation 1.

The majority of the remaining curves in figure 4 are for transistors with a gap in gate to channel coverage near the drain end of the channel. It is obvious that the subthreshold swing of the drain gapped transistors (DGT's) is dependent on the drain voltage. It seems to follow a parabolic dependence on drain voltage until the swing reaches a constant swing magnitude. If the drain voltages at which this first occurs are plotted against the square of the magnitude of the gap size again a linear relationship becomes evident, as illustrated in figure 5. This is simply the motion of the drain depletion region into the channel under the influence of the drain field. The slope of the resulting line can be used to estimate the surface doping concentration using a standard depletion width equation [4].

$$y_d^2 = \frac{2\epsilon_s}{qN_A}(V_{bi} - \Psi_s + V_d) \qquad (1)$$

The predicted swing magnitude calculated using the surface concentration extracted by this method agrees with the observed value.

5. CONCLUSIONS

A novel technique used to manufacture transistors with quantifiable gaps in gate to channel coverage near the source or drain has been described. This technique has a significant advantage over previous methods since transistors with a known gap can be identified and their characteristics measured. These have been used to demonstrate that asymmetries in the subthreshold region of gapped transistors are caused by the requirement for the uncovered channel region to be inverted by the fringing field of the gate. In the SGT case the subthreshold swing is consistently increased and was observed to be linearly dependent on the size of the gap. However the drain voltage dependent swing of the DGT has been shown to be caused by drain depletion widening.

6. FURTHER WORK

This research will continue with consideration of the saturation region. The fabrication of polysilicon gated transistors with smaller step sizes is also planned.

REFERENCES

[1] P.K. Ko et al. 1986 IEDM Tech. Dig. p292.

[2] T.Y. Chan et al. IEEE Elec. Dev. Let. vol. EDL-7 no.1 Jan. 1986 p16.

[3] T.Y. Chan et al. IEEE Elec. Dev. Let. vol. EDL-8 no.6 June. 1987 p269.

[4] S. M. Sze. Physics of Semiconductor Devices (John Wiley & Sons, New York, 2nd Ed. 1981).

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of DEC and SERC. J. A. Serack is also supported by BNR and NSERC of Canada.

The Measurement of Transistor Characteristics Using On-chip Switching for the Connection of Instrumentation

D Ward, A J Walton, J M Robertson

Department of Electrical Engineering,
Kings Buildings,
University of Edinburgh,
Edinburgh, EH9 3JL,
Scotland

This paper investigates the feasibility of using on-chip switching for the instrumentation used to measure transistor characteristics. The effect of the switching transistors on the measurements are evaluated by comparing the SPICE parameters extracted from measurements made via the switching transistors with those derived directly. It is shown that accurate SPICE parameters can be extracted from process control chips with on-chip switching.

1. INTRODUCTION

The measurement of transistors is traditionally performed by a parametric tester using a switching matrix to connect the instrumentation to the appropriate pins. Each transistor may use up to four pads and hence much of the area on the test chip is not occupied by the test devices themselves. Not only are a large number of pads required but as a consequence, a large number of expensive relays are necessary and these significantly reduce the speed of measurement. If the switching matrix can be located on the chip then the pad count can be reduced and switching speeds increased. This approach has been used previously to make yield evaluations with contact chains and transistor arrays when only qualitative measurements were required [1]. This paper will address some of the factors that need to be considered when quantitative measurements are to be performed.

2. SIMULATION OF THE SWITCHING TRANSISTOR

Figure 1 shows one configuration for measuring transistor characteristics. Initial simulation of this circuit with pass transistors as switches for both forcing current and sensing voltage were made using SPICE [2]. These indicated that the use of pass

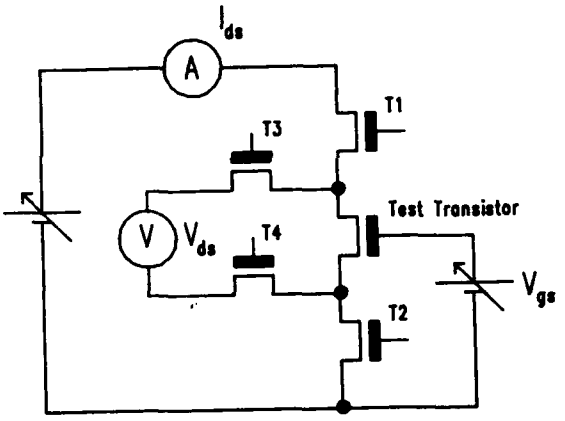


Figure 1. Circuit using pass transistors for on-chip switching to measure transistor characteristics.

transistors in a voltage sensing circuit would cause no appreciable voltage drop across the pass transistor provided the gate voltage on that transistor was at least V_t above the voltage to be sensed. This is illustrated in figure 2. However it was found that in a current forcing situation, the transistor $T1$ in figure 1 limited the current flow in the circuit, saturating at a lower I_{ds} due to its non-zero source voltage. This is shown in the simulation of figure 3 where current limiting can be observed for transistor $T1$. The above

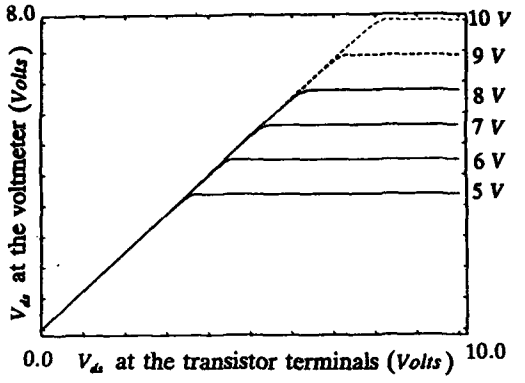


Figure 2. SPICE simulation of V_{ds} sensed by the voltmeter plotted against V_{ds} at the transistor terminals for different values of gate voltage on the pass transistors T3 and T4.

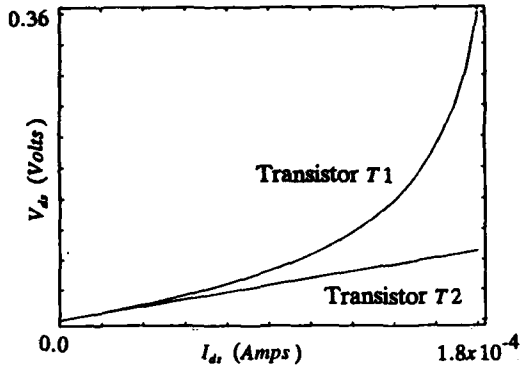


Figure 3. Simulated relationship between V_{ds} and I_{ds} for the transistors T1 and T2 in Figure 1.

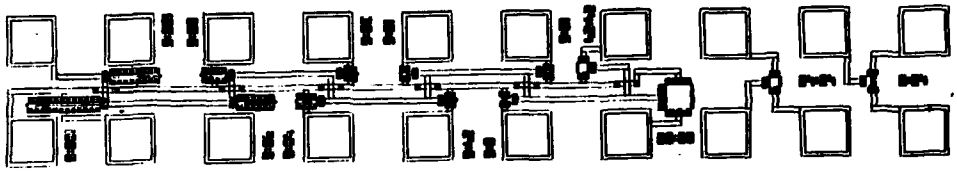


Figure 4. Module used to examine the performance of pass transistors. It contains two test transistors and twelve pass transistors.

simulations were taken into account and two test chips designed to evaluate the performance of the switching transistors. These both contained a number of different measurement configurations, with a range of pass and test transistor geometries. For example, the configuration given in figure 4 allows two test transistors with different geometries to be accessed using one of the twelve pass transistors, all of which have different dimensions. The layout has been designed so that the test transistor characteristics can also be directly measured to allow the switching performance of the pass transistors to be evaluated.

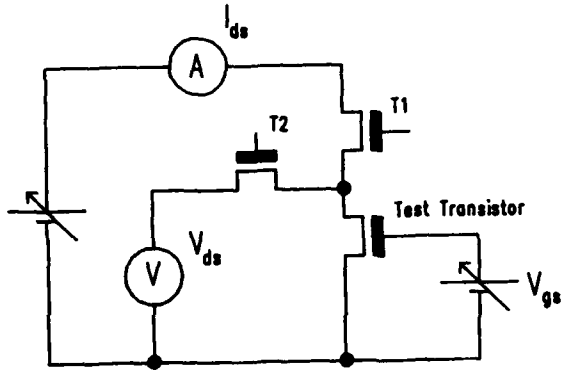


Figure 5. Circuit for measuring transistor characteristics using only two pass transistors.

3. MEASUREMENTS

The simplest design that includes both current forcing and voltage sensing pass transistors is shown in figure 5. It will be used in the following measurements to highlight the problems which need to be considered. Figure 6 shows the characteristics of a transistor measured using pass transistors with different dimensions and compares them with those made directly on the test transistor. It can be observed that the transistor characteristics in the linear region are less sensitive to the dimensions of the pass transistors since $T1$ does not limit the current.

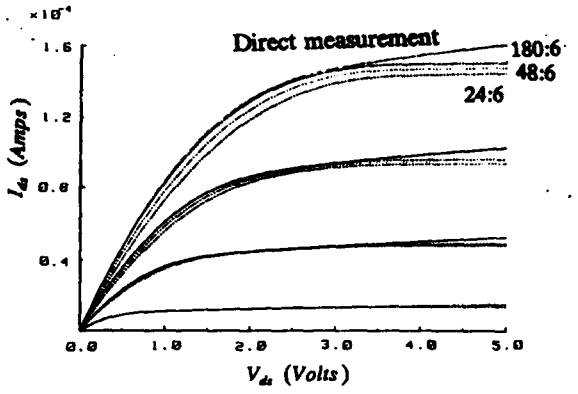


Figure 6. The effect of different sized pass transistors (T1) on the measured I_{ds}: V_{ds} characteristics for the test transistor, compared with the direct measurements.

However, in the saturated region this is not the case and a larger pass transistor helps to obtain a better representation of the device under test. Increasing the size of the pass transistors *ad infinitum* is not a practical solution to the measurement. The current through the test device is limited because V_{gs} of the pass transistor never reaches $5V$ because there will always be a voltage drop across the transistor being measured. Figure 7 shows how the characteristics obtained with higher gate voltages on the pass transistor overcome this problem with the measurements approaching the true characteristics as V_{gs} is increased.

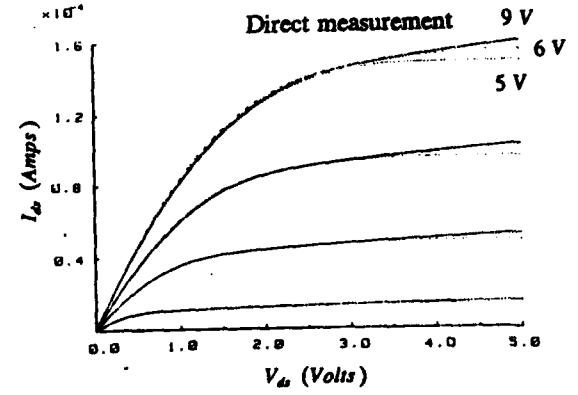


Figure 7. The effect of increasing the gate voltage of the pass transistor(T1) on the measured I_{ds}: V_{ds} characteristics for the test transistor, compared with the direct measurements.

4. SPICE MEASUREMENTS

One very sensitive test of the accuracy of the measurements is to extract SPICE parameters from measurements made through the pass transistors and then compare them with those obtained directly from the transistor terminals. PARAMEX [3] was used for these extractions and derives SPICE 2 level 3 parameters sequentially without recourse to any numerical optimisation. As a result all the extracted parameters have physical significance and can consequently be used for process control.

Table 1 compares the individual SPICE parameters extracted from direct measurements on the test transistor and those made through a pass transistor when its gate voltage was set at $9V$. The errors in the characteristics simulated using SPICE parameters derived from measurements via pass transistors compare very favourably with those obtained when the parameters are extracted from measurements made in the normal manner. This is illustrated in table 2 where the SPICE parameters have been extracted from measurements performed for different values of V_{gs} on transistor $T1$.

SPICE parameters	Measured at the transistor terminals	Measured through the pass transistor (120 x 6 μm V_g = 9V)
t_{ox} (m)	8.5×10^{-8}	8.5×10^{-8}
X_j (m)	1.0×10^{-6}	1.0×10^{-6}
N_{fs} (m^{-2})	1.0×10^{15}	2.9×10^{15}
V_{to} (V)	1.03	1.06
γ ($\sqrt{V}$)	0.65	0.62
L_{del} (m)	1.35×10^{-6}	1.35×10^{-6}
ΔW (m)	1.46×10^{-6}	1.47×10^{-6}
μ_o ($m^2V^{-1}s^{-1}$)	0.085	0.087
θ (V^{-1})	0.051	0.067
V_{max} (ms^{-1})	4.28×10^5	5.69×10^5
η	0.159	0.134
δ	0.367	0.329
κ	0.402	0.495

Table 1. SPICE parameters extracted from direct measurements on the test transistor compared with those using measurements via the pass transistors.

V_{gs}	Average % rms error with the measured characteristic				
	Terminal measurement	Pass transistor gate voltage			
		5	6	8	9
2.0	11.4	19.7	18.5	16.2	16.3
3.0	6.3	7.2	8.6	7.2	7.2
4.0	2.5	3.0	4.1	3.0	3.0
5.0	0.6	1.3	2.3	1.4	1.3

Table 2. Rms error for simulated characteristics using SPICE parameters extracted from conventional measurements and via the pass transistors for different values of V_{gs}.

Figure 8 shows the variation of κ as a function of gate voltage on the pass transistor. This parameter has been selected since it is the SPICE parameter which characterises the slope of the $V_{ds} : I_{ds}$ curve in the saturation region. From figure 7 it would be expected that κ would increase rapidly between 5V and 6V and, as the gate voltage is further increased would, approach the value extracted from direct measurement on the transistor. This is the case in figure 8 with the small overshoot being explained by variations in other parameters which offset the slightly increased value of κ.

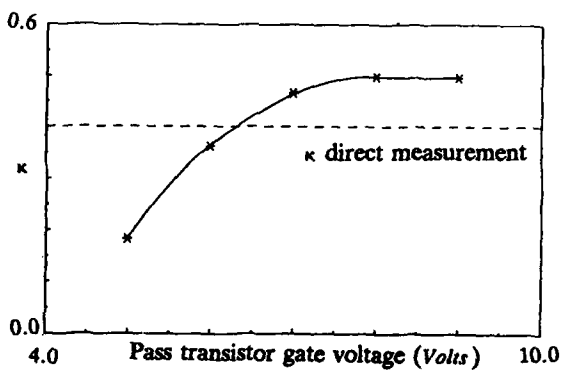


Figure 8. The effect of increasing the gate voltage of the pass transistor(T1) on the measured κ of the test transistor. The dotted line indictes the κ value obtained by direct measurements.

5. CONCLUSIONS

The results which have been presented illustrate that it is feasible to implement a switching matrix on-chip which can be used to measure transistor characteristics. It has been demonstrated that SPICE parameters can be extracted with an accuracy comparable to those derived using conventional techniques. This opens up the possibility of process control chips being measured using equipment with no switching matrix; the individual devices being addressed through pass transistors. With fewer pads required, these test chips can be designed to occupy reduced.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the EMF staff who were involved in the fabrication of the wafers and Dr H Coltman of GEC for useful discussions. D Ward would also like to acknowledge financial support from GEC.

REFERENCES

1. A.J. Walton, J.M. Robertson, R. Holwill, M.B. Moore, Electronics Letters, Vol 21, No 4, pp 127-128, February 1985.
2. A. Vladimirescu, S Liu, "The Simulation of MOS Integrated Circuits Using SPICE2", Memorandum No. UCB/ERL M80/7, University of California, Berkeley, Feb 1980.
3. A. Gribben, J.M. Robertson, A.J. Walton, Semicon, pp 186-202, Birmingham, Sept 86.

A 2D CARRIER PROFILING TECHNIQUE FOR VLSI PLANAR STRUCTURES

C Hill, P J Pearson, B Lewis, A J Holden, R W Allen

Plessey Research Caswell Ltd., Caswell, Towcester, Northants., England, NN12 8EQ

A novel technique for obtaining two-dimensional free carrier profiles with high spatial resolution has been developed. The technique involves fully automated anodic sectioning of resistor structures on a special VLSI test chip and computer analysis of the data. Application of the technique to determination of the 2D distribution of boron in silicon under an oxide mask edge after implant and anneal is described, and results at a spatial resolution of 600Ax600A and sensitivity of 10^{17} ions/cc are presented.

Introduction The need for accurate 2D process models for VLSI has been recognised for many years, and interestingly sophisticated software has become available. There is, however, almost no experimental data at adequate resolution and sensitivity to make these models accurate and reliable for modelling of current one micron geometry VLSI structures. The technique described here has been under development at Caswell for 6 yrs[1] and is designed for a spatial resolution of about 100Ax100A and concentration sensitivity of about 10^{16}ions/cc of dopant.

The principle of the technique is to fabricate identical planar resistor structures, electrically isolated from the substrate and each other, and to monitor the conductivity changes as thin layers of different geometries are removed from each structure. Mathematical analysis has shown that with a minimum of 3 structures, sufficient data is obtained to recover the original conductivity distribution, and hence by use of published mobility data, the ionised dopant distribution. An example of one such analysis is shown in Fig.1. In order to realise this technique experimentally, four essential components have been developed: a special test chip, a fabrication sequence, a fully automatic anodic sectioning and conductance monitoring kit, and a computer analysis programme for conversion of data to 2D profile.

The Test Chip This consists of a 7 layer mask set, 6 of which are shown superimposed in Fig.2. In addition to the basic three identical resistor structures, 16x8 microns in size, (C, G and K), 9 other variants are incorporated, allowing the effects of implant asymmetry to be assessed, and the progress of the anodisation process itself to be monitored. The resistors are contacted by four-terminal connections through substrate diffusions and doped polysilicon rails to metal pads around the periphery of the chip.

Fabrication sequence The mask set is used to fabricate part-processed stock that contains the resistor connections and resistor windows. In subsequent fabrication steps, the desired implant and anneal is incorporated in these windows and the starting geometry for sectioning is defined. One resistor is completely protected so that subsequent anodic sectioning will remove horizontal strips from the complete resistor. The other two resistors are anisotropically plasma-etched, so that the planar portion of the resistor is removed, leaving the portions under the mask edges. The protective oxide mask is then etched from one of these structures.

Automatic Sectioning and Conductance Measurement A computer-controlled anodisation cell and probe card rests on the 5" slice, defining and sealing the chip to be measured. A cyclic

anodisation and etch process, using amyl phosphate and 15:1 buffered HF, removes 200A thick sequential layers from the resistor structures; horizontal (H) planar layers from structure K; vertical (V) planar layers from structure C; L-shaped horizontal and vertical (HV) layers from structure G. Conductance measurements on each structure are made and logged automatically at each step, and when plotted graphically are of the form shown in Fig.1d.

Data Analysis To determine the 2D conductivity distribution in the resistor structures, a region representing a vertical section through the resistor at right angles to the current flow is divided into a finite matrix of rectangular elements. The sequential sections are mapped onto this matrix and different conductances are assigned to each matrix element until a set consistent with all three conductance profiles (H, V, HV) is found. This can be done manually, or by using a computer fitting programme to a generalised 2D exponential function with 12 fitting parameters(Fig.1c), and a topography mapping routine which simulates the sectioning topography with a 6 point linear interpolation routine (Fig.1a and 1b).

2D Boron Concentration Distributions Boron distributions were determined for resistors implanted with $5x10^{14}$ B^+ ions at 30KeV into 2 ohm cm N type [100] silicon, and annealed at 1000°C for 30 minutes in dry nitrogen. After fabrication of the H, HV, and V resistor structures, the sequence of conductance measurement, 100 volt anodisation and etch was repeated until conductance changes in the planar H structure indicated that the p-n junction had been reached (35 cycles). The differential conductance thickness data obtained from the three basic structures is shown in Fig.3. SEM examination of the structures before and after sectioning showed that the topography was very different from that of Fig.1, entirely due to the shallow angle (16°) of the masking oxide edge (Fig.4) produced by the wet-etching characteristics of the photoresist defining layer on the very thick oxide needed to protect against the many subsequent anodisation-etch cycles. Lateral erosion of this oxide during plasma etching removed a portion of the doped silicon and a 60° (rather than vertical) section through the structure. Anodisation caused further lateral erosion, resulting in removal of wedge-shaped (8°) surface sections, in addition to the "vertical" sections expected from the V structure. All these effects were, however, reproducible and well-controlled: the evolution of the topographies is shown in Fig.4.

Because of these large deviations from the original computer model assumptions, manual fitting was used initially to produce a self consistent conductance matrix, based on 600x 600A elements. The resulting concentration distribution is shown in Fig.5, and represents a fit to better than 1% to the original data. An unexpected result from this best fit was the evident loss of boron from the near- surface region, presumably caused by segregation to the mask oxide.

Acknowledgements

This work has been supported by the Alvey Directorate and sponsored from the R.S.R.E. Establishment and the Plessey Company. The permission of both bodies to publish is acknowledged, as is also the invaluable assistance of sandwich students Graham Cooke, Stuart Gurden and Richard Beanland whose enthusiasm and hard work overcame many of the problems encountered in realising this technique.

Reference

1. C Hill and A L Butler ESSDERC/SSSDT 1983 Inst. Phys. Conf. Series No.69 161-180 (1983)

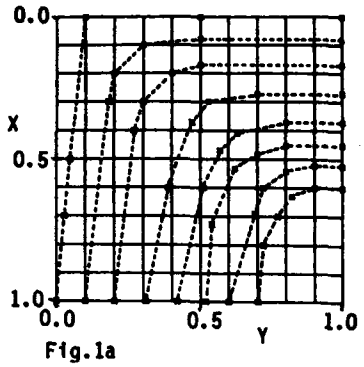


Fig.1a

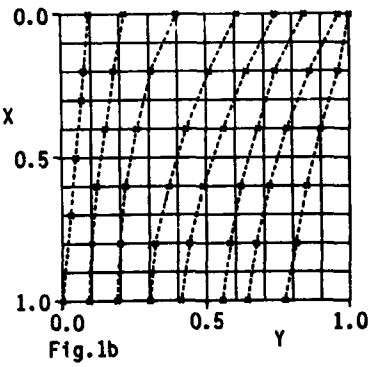


Fig.1b

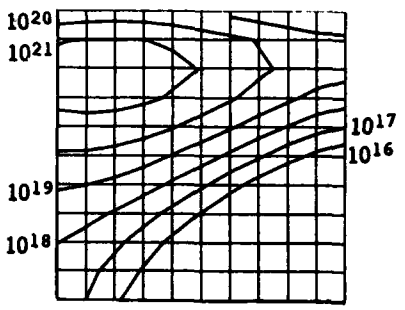


Fig.1c

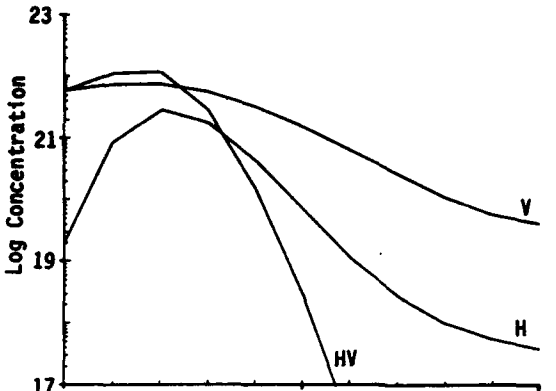
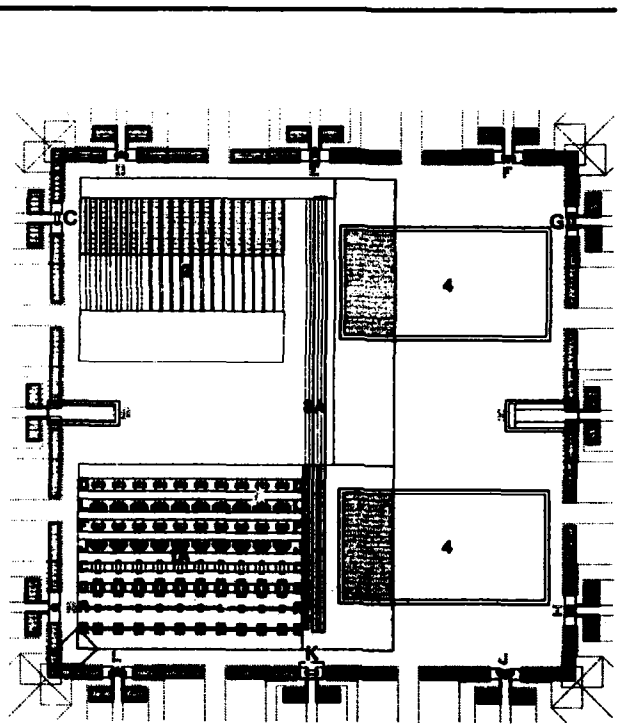


Fig.1d

FIGURE 1

Mathematical simulation of the determination of the 2D free carrier distribution under an implanted and masked edge. (a) and (b), Experimentally measured evolution of surface topography in anodically sectioned HV and V structures respectively; (c) Simulated 2D concentration contours under a mask edge; (d) Simulated conductance - depth data for the three basic resistor structures, sectioned according to the topographies shown in (a) and (b) and containing the simulated dopant distribution shown in (c). The computer fitting programme took the data from (a), (b) and (d) and generated a 2D concentration contour plot indistinguishable from (c).



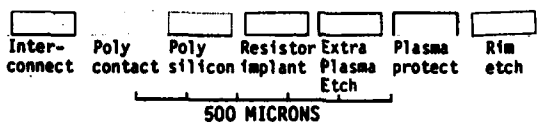


FIGURE 2

Central area of the 2D test mask showing the 12 resistor structures around the periphery labelled A-L.

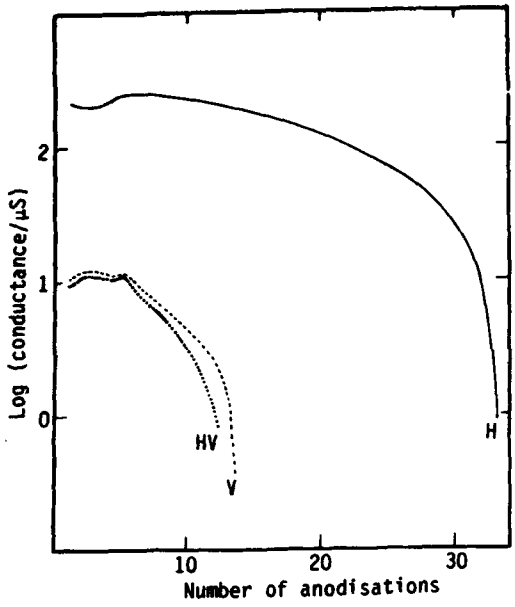


FIGURE 3

Experimental conductance-depth profiles measured during sequential anodic removals of boron implanted and annealed silicon layers from 'vertical' sections (V), horizontal and 'vertical' sections (HV) and horizontal sections (H) of resistor structures C, G and K respectively.

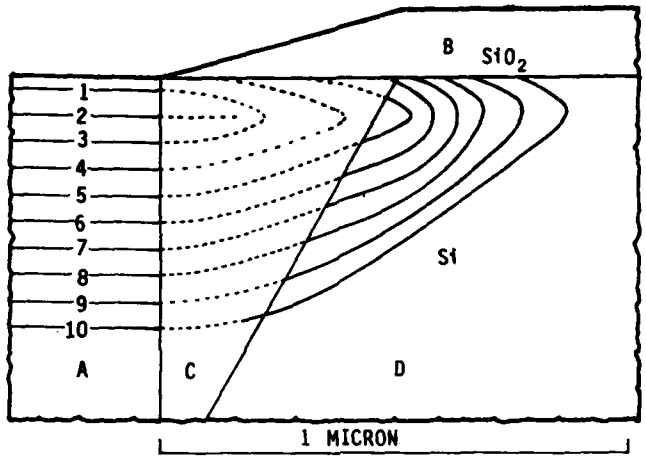


FIGURE 5

Experimentally determined 2D boron concentration contours in a section through the edge of an implanted and annealed resistor structure (for details see text), derived from the data in Figs.3 and 4. Regions of the structure are:- A, planar portion of concentration profile; B, masking oxide at implant stage; C, silicon lost during plasma etch stage; D, silicon remaining at start of anodic sectioning stage. Concentration represented by contours in ions/cc are:-
1, 8.3E18; 2, 9.5E18; 3, 8.8E18; 4, 7.5E18; 5, 5.7E18; 6, 4.1E18; 7, 2.65E18; 8, 1.6E18; 9, 6.2E17; 10, 1.5E17

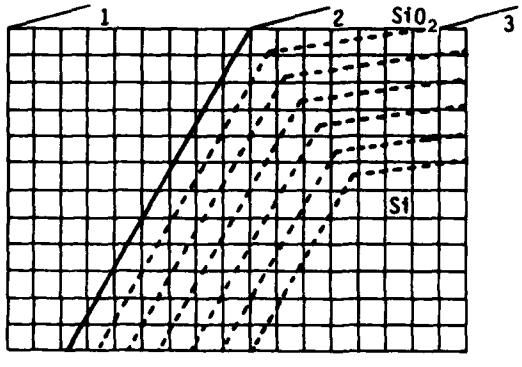


Fig.4a

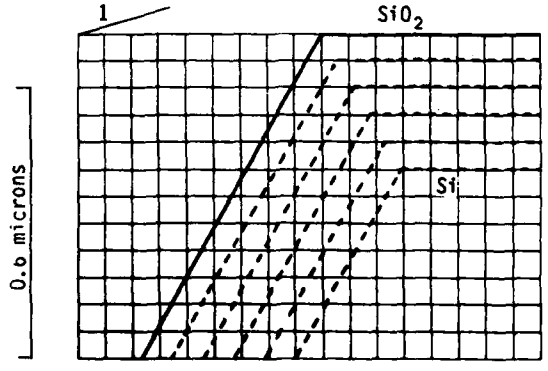


Fig.4b

FIGURE 4

Experimental evolution of surface topography after plasma etching and anodic sectioning of (a) HV structure and (b) V structure, showing the position of the masking oxide edge after: 1. implant and anneal, 2. substrate plasma etch and 3. completion of the first anodisation-measurement-etch cycle.

STATISTICAL DEFECTIVITY CONTROL FOR VLSI DEVICES

R. Traversini, A. De Lisio, M. Tosi and G. Barbuscia

SGS Microelettronica, Via C. Olivetti, 2
20041 Agrate Brianza (MI) - Italy *

A method used in designing test chips for defect monitoring of VLSI processes is presented. The purpose of the method is to maximize the accuracy of defect density estimations. This is achieved by appropriate sizing of test chip structures. The optimal dimension of test structure is computed using a model relating the confidence interval for the defect density D with the sample size and test structure dimension. The Poisson and negative binomial yield models are used in the calculation. The application to real life VLSI processes is outlined.

1 Introduction

Point defects are the major cause of yield losses in VLSI manufacturing and hence a major concern in VLSI devices design.

Defects are responsible for localized modifications of the device structure, such as inter–layers shorts, and may lead to electrical failure of the chip.

The probability for a point defect to produce a failure depends on the its size, on the local device structure and on lay–out rules [1].

Monitoring and improving product yield require tools for detecting and discriminating failures of different kind. This is generally accomplished processing wafers containing test chips with appropriate structures for detecting particular types of failure. (As an example interconnect snake/comb structures are used to investigate interconnect shorts and opens.)

Electrical failure statistics are then collected from test chip testing and subsequentely elaborated to extract killer defect densities.

The elaboration is based upon yield models relating yield figures to defects distribution and device geometry [2].

By means of standard statistical process control tools (such as control charts), the resulting estimates of defect densities are compared to target or to trend limits in order to detect out–of–control situations or to verify the results of yield improvement efforts.

Since the use of test wafers is a costly and time consuming procedure it is mandatory to maximize the accuracy of each single estimate of defect density.

The aim of this work is to reach the appropriate sizing of test chip structures which renders defect estimates as much accurate as possible.

2 Test chip structure sizing

In the following the true defect density will be indicated with D whereas its estimation D^*.

We will assume that the number of failure occurring on a test structure is directly proportional to a dimensional feature of the structure, A, such that the product AD is equal to the average number of failure detected on the structure. This last statement actually is the operational definition of D.

A simple framework for the defect estimation extraction is considered. We assume D^* values are obtained averaging at lot level the estimated defect density D^*_w of each wafer. This in turn is derived from the

* This work has been partially supported by CNR, Progetto Finalizzato "Materiali e Dispositivi per l'Elettronica a Stato Solido"

measured yield Y_w^* of the related test chip structure in the wafer using some yield model.

In this work we are concerned with defect control and yield diagnosis; in this context D^* is used for comparison with target values and control limits to check the process status.

The point here is in minimizing the statistical variability of the estimate D^*. This variability depends on the sample size (the number n_w of tested wafers and the number n_d of test chip in each wafer) and on the structure size A.

D^* uncertainty can be described in term of confidence intervals. A confidence interval for D is an interval containing the estimate value of defect density and such that the true value of D has a predefined probability to belong to it. It will be denoted by (D_l^c, D_u^c), where c is the so called confidence coefficient and is such that:

$$Prob(D \in (D_l^c, D_u^c)) = 1 - c$$

The problem can be stated in the following way: find the value of A which correspond to the minimum value of the amplitude $(D_u^c - D_l^c)$ of the D confidence interval, given the values of n_d, n_w and c.

This problem will be discussed in the following section. First we present a solution based on the uniform Poisson yield model, then the more general case of Stapper negative binomial yield statistics is studied.

2.1 Poisson yield model

In the Poisson yield model the defect generating process is modeled as a Poisson process acting in the same way on the whole population of chips in the lot.

The actual behaviour of defect yield losses is generally different from the simple Poisson case. More complex models are required to describe the observed distribution of defect density. But this model is appropriate for a qualitative description of the D^* confidence interval as function of A.

The probability $P(x)$ to collect x defects on a test structure is:

$$P(x) = \frac{e^{-\mu}\mu^x}{x!}$$

where μ is the average number of defects on the test structure.

From the definition of D given in the previous paragraph, in this case $\mu = AD$. The yield Y is the probability having no defect and then $Y = P(0) = e^{-\mu}$.

Combining these relations, one can estimate the wafer defect density from the formula $D_w^* = -ln(Y_w^*)/A$, where Y_w^* is the wafer yield. Then, the average defect density estimation is $D^* = (\sum_{w=1}^{n_w} D_w^*)/n_w$.

Let us define m as the number of failed test structures in a wafer. In the Poisson model, m is a binomial random variable with $N = n_d$ and $p = 1 - e^{-AD}$. A confidence interval (p_l^c, p_u^c) for the parameter p can be obtained solving for p_u^c and p_l^c the system:

$$\begin{cases} \sum_{j=0}^{M} \binom{n_d n_w}{j} (p_u^c)^j (1-p_u^c)^{n_d n_w - j} = c_1 \\ \sum_{j=M}^{n_d n_w} \binom{n_d n_w}{j} (p_l^c)^j (1-p_l^c)^{n_d n_w - j} = c_2 \end{cases}$$

where $M = n_d \sum_{w=1}^{n_w}(1 - Y_w^*)$ and $c_1 + c_2 = c$.

From (p_l^c, p_u^c) a confidence interval for D is calculated using the relation $D = -ln(1-p)/A$.

In figure 1 the confidence interval amplitude $\Delta = (D_u^c - D_l^c)$ versus structure dimension A is shown for different values of defect density D, as computed using the prevoius formulas. The value of M is assumed equal to $n_d n_w(1 - e^{-AD})$ and $c_1 = c_2 = c/2$.

The unit of A is arbitrary, as implied by the definition of this quantity: if the structure scaling dimension is an area, one can choose cm^2 as dimension unit; as a consequence, D and Δ on the vertical axis are measured in defects/cm^2. The same plot can be used with linear dimensional unit or for contact defect, whose test structures are measured in terms of number of tested contacts. The plot relates to fixed sample size ($n_d = 35$, $n_w = 25$) and a value of 0.05 is assumed for the confidence coefficient.

The plots in figure 1 show the essential trends of confidence interval amplitude. For given D, c and sample size, the amplitude Δ is large for small values of A and , as A increases, it becomes smaller achieving a minimum for a value of A depending on the parameters of the plot. For larger values of A the confidence interval amplitude increases.

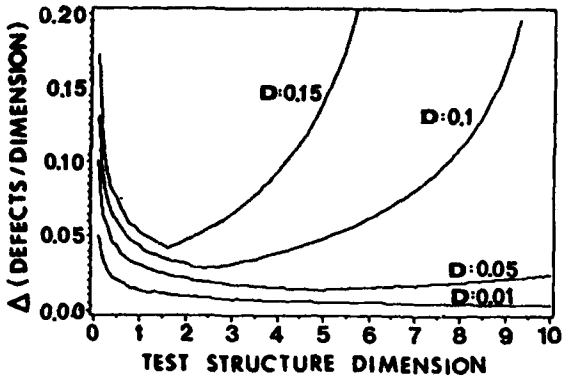


FIGURE 1

2.2 Negative binomial yield model

The same problem can be solved in the case of Stapper's negative-binomial yield models. These are based on the assumption of a non-uniform Poisson process for defect generation. The true defect densities is distributed on the wafer (or at some other aggregation level) according to the Gamma probability density function:

$$f_D(z) = \frac{z^{\alpha-1}e^{-\frac{z}{\beta}}}{\beta^\alpha \Gamma(\alpha)}$$

The true defect density D is the expected value of this distribution and it turns out to be $D = \alpha\beta$. Then, the estimation of D is based on the estimation of the product of the two parameters of the Gamma density function [3].

In solving this problem we turn to Monte Carlo simulation. A SAS based procedure has been built up [4] which simulates the defect generation process on lots composed of several wafers containing test structures of given size. The number of defects located on a single structure is extracted from a Poisson distribution whose parameter is extracted, for each wafer, from a Gamma distribution.

This is done for a number of lots. For each lot the average yield and the yield variance are computed. In the framework of the Stapper model one can easily show that:

$$\begin{cases} E(Y) = \frac{1}{(1+A\beta)^\alpha} \\ Var(Y) = \frac{1}{(1+2A\beta)^\alpha} - E^2(Y) \end{cases} \quad (1)$$

The first of these formulas is the well known relationship between the espected yield $E(Y)$ and A, α and β in the Stapper case [2]. The second is obtained from the relation $Var(Y) = E(Y^2) - E^2(Y)$ and evaluating the integral:

$$E(Y^2) = \int_0^\infty (e^{-Az})^2 f_D(z)dz$$

The average yield and yield variance of each lot are used in solving the equations (1) for α and β. A lot estimation of β is obtained solving by iteration the problem:

$$x = \phi(x)$$

where:

$$\phi(x) = (1-2x)^\gamma - 1$$
$$\gamma = ln(E(Y))/[ln(E^2(Y) + Var(Y))]$$
$$x = A\beta$$

The corresponding value of α is then computed from any of the equations (1) and $D^* = \alpha\beta$ may be obtained for the estimate of D for the particular lot.

The statistics of D^* is cumulated and when an appropriate sample has been collected the confidence interval of smallest amplitude with confidence coefficient equal to c is computed. The calculation is then repeated for different values of A.

In figure 2 one curve obtained this way, corresponding to $\beta = 0.1$, $\alpha = 1$, $n_d = 35$, $n_w = 25$ and $c = 0.05$, is shown. The true value of the average defect density id $D = \alpha\beta = 0.1$ defect/(dimension unit).

The trend of confidence interval amplitude as function of A is the same as in the Poisson case. The minima of Δ for different values of D correspond to the same values of A obtained in the Poisson model. In the negative-binomial model the value of Δ in the same conditions of sample size and structure dimension is larger than in the Poisson case, as expected because of the spread of D described by the Gamma density distribution function.

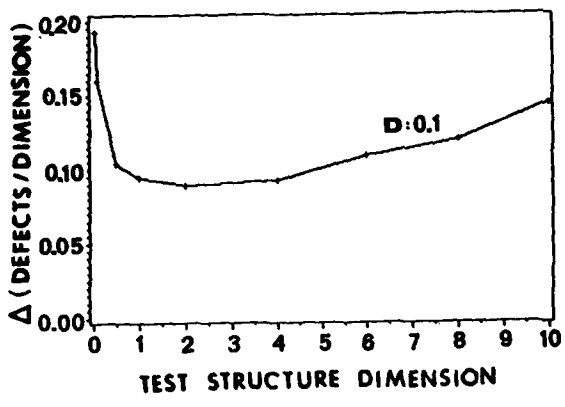


FIGURE 2

3 Results and applications

The plots in figure 1 and 2 can be used to determine the optimal structure sizing and to estimate the relative improvement in accuracy resulting from the appropriate choice of A.

For comparatively high defect densities ($D \simeq 0.1$ defect / (dimension unit) or greater), choosing structure dimensions around $2 \div 3$ dimension units results in an improvement of test chip structure accuracy, whose extent depends on the defect statistics.

On the other hand, in low defect density situation, there isn't a minimum for Δ in the range of pratical values for A. The curve steepens for $A \leq 1$, while for larger values of A it tends to a nearly constant level. Due to limitation in chip size, the biggest dimension compatible with chip area limits is required to achieve minimal error.

The method of optimal sizing of test chip structure outlined in this work has been used in SGS in the design of a defect test chip for the 1Mbit CMOS EPROM device developed in SGS.

The defect test chip contains a set of structure designed for detecting failures in the critical process layers (Active Area, two Poly layers, Metal and contacts). Oxide layers are not considered because they are monitored routinely by the means of specific test chip, on which however the same sizing procedure has been applied.

The defect density level to be monitored at each layer spans between the present level of D and the target level as deduced from the target yield of the device.

Each test structure has been sized for the lowest defect density level (the target level) and then intermediate connections have been established to ensure the availability of test zones of smaller area, which are more appropriate when the defect density is still higher.

Given a target defect density ≥ 0.01 defect/(dimension unit), the optimal sizing is achieved for large values of A. However for $A > 3 \div 4$ no pratical improvement can be gained increasing the structure dimension. In this situation the best choice is between these last figures and the layer dimension of the device layout. This latter can be estimated using standard layout extraction tools like the MASKAP program.

When the device layer critical dimension is around the above limiting value for A is useful to choose it as the test structure dimension. Indeed, this choice allows direct comparison of test chip electrical yield with device yield with minimum scaling factor.

Intermediate zone sizing is driven by the smallest zone to be defined, which in turn is determined by the highest defect density to be detected. If a starting value of about 0.1 defect density unit for D is assumed, this size is found to lie around $1 \div 2$ dimension units.

References

[1] Stapper C. H., IBM J. Res. Develop., Vol. 27, 549 (1983)

[2] Stapper C. H., IBM J. Res. Develop, May 1976, 228 (1976)

[3] Winter C. L. and W. L. Cook, IEEE Journal of Solid-State Circuits, SC-21, 590 (1986)

[4] *SAS User's Guide: Basics and Statistics*, SAS Institute Inc., (1985)

IP.13. - INVITED PAPER

THE PHYSICS OF SILICIDE BASE TRANSISTORS

E. Rosencher, F. Arnaud d'Avitaya, P.A. Badoz, G. Glastre, and G. Vincent
Centre National d'Etudes des Télécommunications - BP : 98 - Chemin du Vieux Chêne - 38243 MEYLAN CEDEX - FRANCE

Epitaxial Si/$CoSi_2$/Si structures can be grown under ultra-high vacuum conditions. The metallic $CoSi_2$ films can be extremely thin typically between 1 nm and 20 nm. The electrical properties of these heterostructures are presented, mainly the transport of electrons in the metallic films parallel to the interfaces and the transfer of electrons through the metal film. The influence of pinholes in the $CoSi_2$ layers will be discussed.

I. Introduction

Thanks to advances in ultra-high vacuum technology, it has recently become possible to realize epitaxial Semiconductor / Metal/ Semiconductor (SMS) structures /1, 2/ using a Si/$CoSi_2$/Si sandwich. The rather small lattice mismatch (~ 1.2 %) between Si and $CoSi_2$ crystals as well as their similar cubic structures allow the production of monocrystalline Si/$CoSi_2$ and Si/$CoSi_2$/Si heterostructures. These SMS structures open the way to promising ultra-low base resistance devices for millimeter wave applications.

In this paper, we intend to review the main results on the physics of metal base transistors (MBT). This latter denomination adresses two different kinds of structures : the SMS-Transistor where the $CoSi_2$ film is intended to be continuous and Permeable Base Transistors (PBT) where discontinuities in the metallic film are intentionally introduced by nanolithography techniques. Though this paper is mainly focused on the electrical properties of these structures, some informations on the morphological quality of these sandwiches are given in Sec. II, which are necessary for the understanding of the transport properties. In Sec. III, parallel transport in the ultra-thin metal films will be addressed while the perpendicular transport through the SMS structure will be developed in Sec. IV. We shall describe the current status of the Si/$CoSi_2$/Si permeable base transistor in Sec. V. A brief discussion is given in Sec. VI which tentatively describes the general trends in the future research on metal base transistors.

II. Morphology of Si/$CoSi_2$/Si heterostructures

This section is intended to describe the results necessary as a background for the understanding of transport properties rather than to provide an explanation to the growth mechanisms, which is still clearly lacking. The morphology observations are made using Scanning Electron Microscopy (SEM), Transmission Electron Microscopy (TEM) and High Resolution Transmission Electron Microscopy (HRTEM).

One of the key points for the growth mechanisms is the cleaning of the silicon surface before $CoSi_2$ formation. Up to these days, we have used the Shiraki process, the details of which are described in Ref. /3/ and /4/. This process is based on chemical oxidation-etching cycles, leaving a very thin protective SiO_x film which evaporates when heated above 750°C. The SMS structures are grown in a typical ultra-high vacuum system

described in Ref. /4/. The surface crystallinity and cleanliness are checked in situ by Low Energy Electron Diffraction (LEED) and Auger Spectroscopy (AES). Let us note that no electronic grade materials can be grown when Carbon contamination higher than 5 % of a monolayer is present at the silicon surface /5/.

In our structures, $CoSi_2$ is grown by Solid Phase Epitaxy on <111> Si surface. Co layers are electron gun evaporated under a pressure less than 5 x 10^{-8} Torr with the sample kept at room temperature and then annealed at 650°C (± 25°C) during 10 minutes in order to obtain the $CoSi_2$ layers. Other growth techniques (molecular beam epitaxy, hot substrates, higher annealing temperatures, etc.) have led to lower quality structures /4-7/. Grazing angle SEM and TEM observations have shown that *smooth* $CoSi_2$ layers could thus be grown for a thickness up to 20 nm. Above this limit, the metal layers are rough and discontinuous, leading to a milky aspect of the wafers /4/.

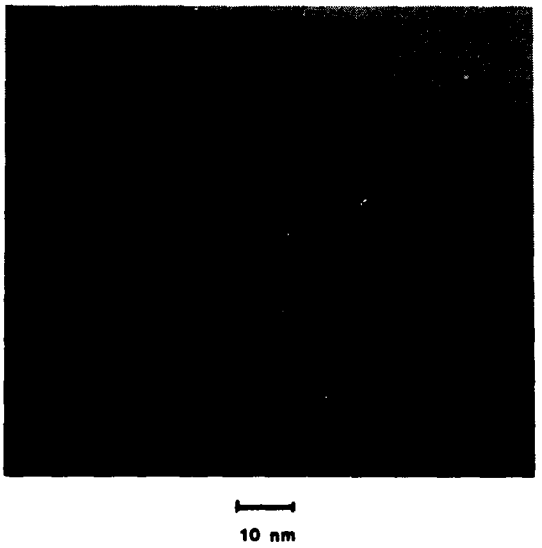


Fig. 1 Cross-section TEM image of a 5 nm thick $CoSi_2$ layer on Si substrate. One notes the bowl-shaped $CoSi_2$ intrusion in Si bulk for an otherwise sharp $CoSi_2$/Si interface.

Cross sectional TEM observations show that, for a thickness less than 20 nm, the interface between Si and $CoSi_2$ is rather smooth, with eventual presence of bowl-shaped $CoSi_2$ intrusions in Si, while the $CoSi_2$/vacuum interface is much smoother (Figure 1) /8/. Plane view TEM photographs of a $CoSi_2$/Si structure are shown in Figure 2a and b. The shape of the Moiré fringes indicates than strain fields are present in the $CoSi_2$ layers and reflects the 1.2 % lattice mismatch between Si and $CoSi_2$. Moreover, no pinholes are observable over few square micrometers /9/.

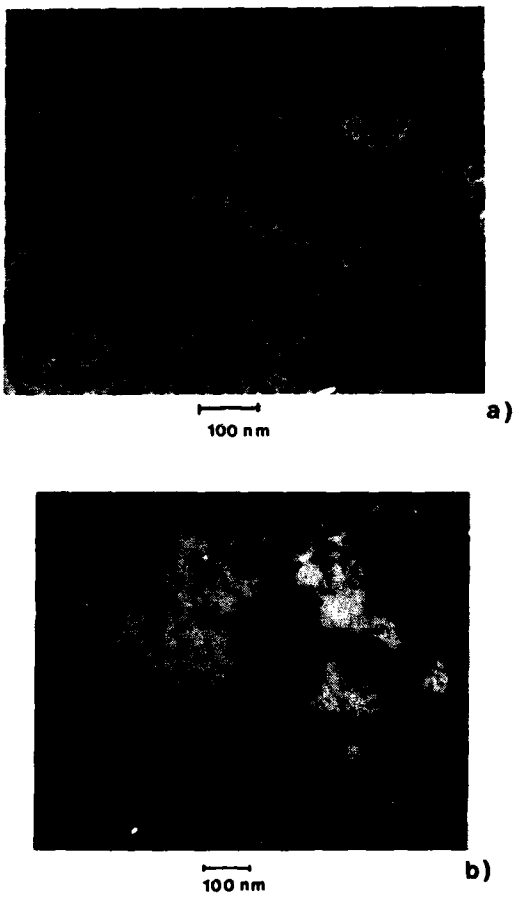


Fig. 2 TEM plane view observations of 5 nm thick $CoSi_2$ layers (a). The bright field image evidences the Moiré fringes (due to lattice mismatch and strains) and the absence of pinholes (b). The dark field image using a diffraction spot specific of type B grains indicates an equivalent density of both grains /9/.

These pinholes, when they are present, appear as regions where Moiré fringes are absent, as described in Ref. /10/. Figure 2b is a plane view TEM observation which stresses the orientation of the $CoSi_2$ crystal relatively to the Si one in the <111> direction. It is clear that $CoSi_2$ is a mixture of two types of grains : type A grains where the Co planes are an extension of the Si bulk planes and type B grains where the $CoSi_2$ lattice is rotated 180° around the <111> direction relatively to the Si lattice /8/. We must stress that some other groups /7, 10/ have found a great majority of type B grains in their $CoSi_2$ films (> 95 %) : the difference in the results of those different groups is still unexplained, though the influence of substrate cleaning is most probably playing a major role.

Silicon layers are deposited by MBE on $CoSi_2$ films at 650°C and doped by Sb coevaporation under a pressure in the 10^{-9} Torr range. Figure 3 is a cross sectional TEM photograph of a $Si/CoSi_2/Si$ heterostructure. It is clear that the roughness of the interfaces has dramatically decreased, the transition between $CoSi_2$ and Si material occurring within one monolayer over long distances. This effect (we call it "planarisation") is still unexplained. Moreover, grazing angle SEM observations indicate that the growth mechanism of Si over $CoSi_2$ is three-dimensional (Fig. 4). This results in a high density of sub-grain boundaries, twins... in the epitaxial Si layer. Finally, let us note that HRTEM observations indicate that, for such small $CoSi_2$ film thicknesses (< 20 nm), the $CoSi_2$ lattice is entirely strained in the $Si/CoSi_2/Si$ sandwich in order to fit the Si bulk lattice /8/.

III. Parallel transport in $Si/CoSi_2$ heterostructures

Because of their outstanding crystalline quality, $CoSi_2$ films are ideal candidates for the study of electron transport in ultra-thin metallic films. Indeed, other possible systems (like Au on NaCl) are usually unstable against exposure to air and temperature cycles.

Hensel et al were the first to show that down to very low thickness, i.e. 10 nm, the film resistivity exhibits little dependence on the $CoSi_2$ film thickness /12/. These thicknesses are much less than the bulk transport scattering length of ~ 100 nm as determined by magneto-resistance measurements, so that boundary scattering of the carriers is essentially specular in this system. In order to account for this phenomenon, these authors have used the well-known Fuchs-Sondheimer theory, introducing a specularity parameter p which is the fraction of electrons specularly reflected from the interfaces. The resistivity ρ of a film of thickness d is thus given by :

$$\rho = \rho_\infty \left[1 - 3/2k \int_0^1 f(u)\, du \right]^{-1}$$

with :

$$f(u) = \frac{(u - u^3)(1 - p)(1 - \exp(-k/u))}{1 - p \exp(-k/u)}$$

where ρ_∞ is the bulk resistivity and k the ratio of the film thickness d to the mean free path λ_e, i.e. $k = d/\lambda_e$. Figure 5a shows the set of curves in $CoSi_2$ films for specularity parameter p ranging from 0 to 1 (i.e., from purely diffuse scattering to purely specular) compared to experimental data from Badoz et al /13/. It is clear that for film thicknesses lower than 10 nm, the variation of the film resistivity with thickness is far steeper than expected from Fuchs-Sondheimer theory.

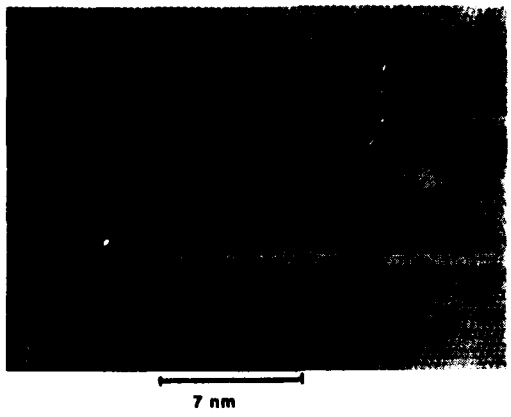


Fig. 3 Cross-section TEM image of $aSi/CoSi_2/Si$ heterostructure. Note the planarisation of interface relatively to Figure 1 /8/.



Fig. 4 SEM image of 23 nm thick Si epilayer on top of $CoSi_2$ (8 nm)/Si_{bulk} structure /11/.

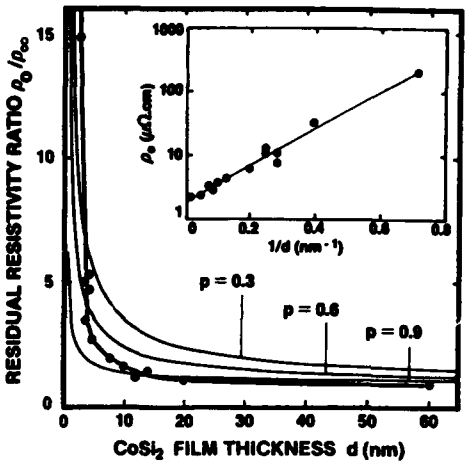
Figure 5b shows the values of the superconducting critical temperature T_c of $CoSi_2$ films as a function of thickness.

Here again, there is an abrupt drop of T_c in $CoSi_2$ films thinner than 10 nm. It has to be noted that these results are different in nature from the usual experiments in thin metal films near the localisation regime ($R \sim h/e^2 \sim 4000\ \Omega_\square$). The sudden change in T_c and ρ occurs in films with resistances in the 30 $\Omega_\square$ range.

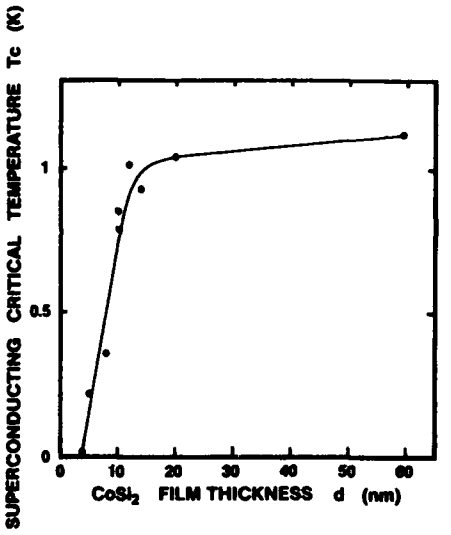
A phenomenological explanation of our results is the presence of a layer of about 5 nm at the $Si/CoSi_2$ interface in which the electronic transport properties are dramatically perturbed : in this region, the carrier mobility is extremely low and no superconductivity occurs. This layer is without doubt $CoSi_2$, as evidenced by LEED, AES, TEM and X-ray photoemission spectroscopy measurements as well as by Hall effect (constant density of carriers in this layer). A possible origin for this perturbed layer

could be the presence of magnetic Cobalt atoms : magnetic impurities are indeed known to be highly effective in quenching superconductivity as well as being efficient scattering centers for electronic transport.

These magnetic Co atoms could stem from either *i)* a small departure from $CoSi_2$ stoichiometry /14/ or from ii) ill-coordinated interface Co atoms as evidenced by careful observation of HRTEM photographs /8/.



a)



b)

Fig. 5 $CoSi_2$ film resistivity ρ_0 measured at 4 K (a) and superconductivity critical temperature T_c (b) as a function of film thickness d /13/.

In an attempt to observe quantization of the metallic electron gas in ultra-thin $CoSi_2$ film, Tunneling Spectroscopy (TS) has been performed in degenerate $Si/CoSi_2$ Schottky diodes (15). Sharp features have been observed in TS spectra up to high energies (600 meV). However, those peaks are weakly dependent on the film thickness (see Figure 6). Very recent experiments tend to prove that those features are due to phonon emission by

the relaxation of hot electrons in the depletion layer of Si. The lack of observable quantization may be due to the high value of the electron energy at the Fermi level E_F, leading to a high value of the quantization energy uncertainty ΔE_F near the Fermi level given by :

$$\Delta E_F/E_F \sim 2\,\Delta d/d \qquad (2)$$

where Δd is the residual film roughness. For a ratio $\Delta d/d$ as low as 1 %, the energy spreading is already in the 100 meV range !

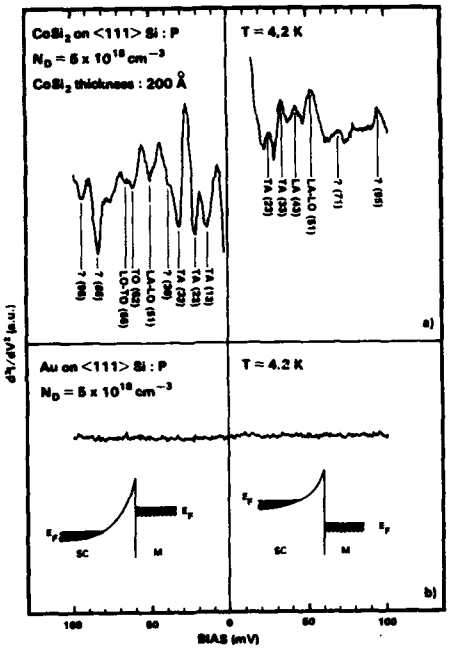


Fig. 6 Tunneling spectra of epitaxial $CoSi_2$/Si tunnel junction (a) and of non epitaxial Au/Si junction (b). The main Si phonons are indicated by arrows /15/.

IV. Perpendicular transport in Si/$CoSi_2$/Si heterostructures

The first evidence of transistor effect in a monolithic SMS structure has been given by Rosencher et al in 1984 /16/. Figure 7 shows the energy band diagram of a SMS transistor. This device consists basically in two back to back Schottky diodes. One of these diodes, the emitter, is forward biased, while the other one, the collector, is reverse biased. The carrier transport between the emitter and the collector is the subject of intensive investigations /16-19/. Two mechanisms may indeed be involved, with relative weight strongly dependent on technology.

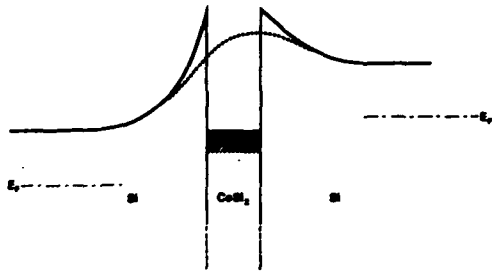


Fig. 7 Schematic energy-band diagram of a SMS transistor in a semiconductor -metal -semiconductor junction (solid line) and in a semiconductor pinhole channel (dashed line).

1. Electrons are emitted via thermionic emission from the forward biased emitter junction into the metal. A fraction of those carriers crosses the metal film via ballistic transport and are collected by the reverse biased collector junction /16, 19/. This mechanism is described by the solid lines in Figure 7.
2. Pinholes are present in the metal film, in which silicon channels are imbedded. The electrons are transferred from the emitter to the collector via those semiconducting channels and the current flow is controlled by the barrier lowering in the pinholes (dashed line in Figure 7) /17, 18/.

A theoretical model has been developed in Ref. /20/ in order to evaluate the different weights of mechanims 1 and 2. The conclusion is that, for usual doping levels, a single 150 nm radius pinhole in the metal base is enough to short circuit the whole ballistic transport in a 20 µm x 20 µm SMS transistor !

It is thus clear that TEM observations, which investigate only few square micrometers of a device, have no statistical significance in order to conclude to the predominance of one mechanism over the other. We have thus

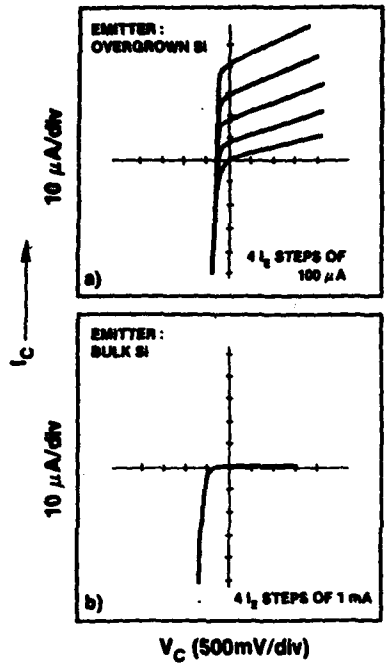


Fig. 8 Common base current-voltage characteristics of a SMS-T with a 7 nm thick base using either the regrown Si(a) or the bulk Si(b) as emitter. The measurements are performed at room temperature.

developed an electrical measurement, a transconductance technique described in Ref. /17/ and /19/, which allows to measure the relative weights of mechanisms 1 and 2. This technique, based on the screening of the collector potential by the metallic $CoSi_2$ film when no pinholes are present, ensures that, in "pinhole-free" SMS, electron transport occurs almost entirely through the metal base. An independent evidence of the dominant role of hot electron transfer through the base in SMS-T is given in Figures 8a and 8b showing the common-base characteristics of a SMS structure (7 nm base thickness) using either the overgrown (Figure 8a) or bulk (Figure 8b) silicon as the emitter. The transfer ratios α are 15 % and less than 10^{-4}, respectively, for the same values of emitter current I_E (500 μA). Since the pinhole current must be of the same order of magnitude in both directions, the value 15 % is clearly due to the transfer through the metal base. The asymmetry of current gain is consistent with the already reported systematic difference in barrier heights Φ_{ms} between the Si bulk/$CoSi_2$ (Φ_{ms} ~ 0.63 eV) and $Si_{epi}/CoSi_2$ (Φ_{ms} ~ 0.69 eV) junctions /19/ : the injected electrons are well above the collector barrier when emitted from the overgrown Si and below when emitted from the bulk Si.

The ballistic transport /21/ in "pinhole-free" SMS-T is described by a mean free path $\lambda_B(T)$, so that the emitter to collector transfer ratio of electrons is expected to be :

$$\alpha = \alpha_o (T) \cdot \exp (- d/\lambda_B (T)) \qquad (3)$$

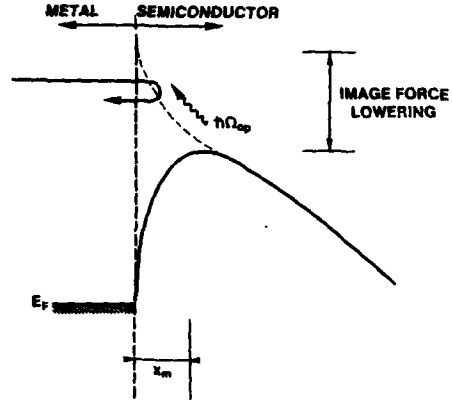


Fig. 9 Electron potential energy versus distance from the collector metal-semiconductor interface. An example of electron backscattering event is symbolized.

where T is the measurement temperature and α_o is the current gain extrapolated to zero metal base thickness. The departure of α_o from unity is due to collector as well as emitter losses :

$$\alpha_o = \alpha_c \cdot \alpha_q \cdot \alpha_e \qquad (4)$$

where α_c is the current gain upper limit associated to scattering in the Si collector, α_q is the quantum mechanical transmission of the base-collector potential barrier and α_e is the emitter efficiency coefficient /21/. The collector scattering contribution is expected to follow :

$$\alpha_c = \exp(-x_m/\lambda_{ph}) \qquad (5)$$

where x_m is the position of the maximum of the collector barrier potential in the image force approximation and λ_{ph} is the mean free path in the Si collector (see Figure 9).

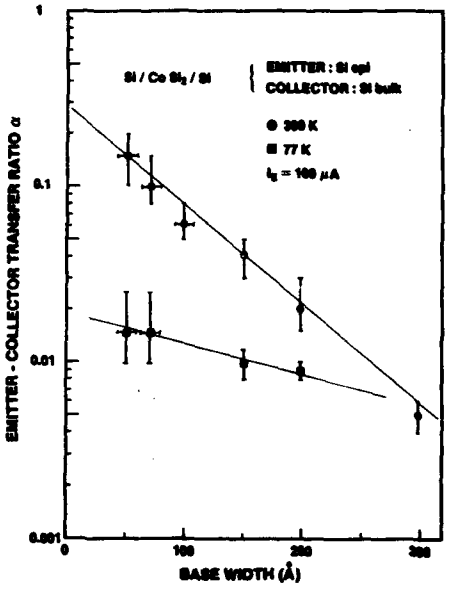


Fig. 10 Transfer ratio versus $CoSi_2$ base thickness measured at 77 K and 300 K in SMS transistor /9/.

Figures 10 and 11 compare the above theory with experiment. Figure 10 shows the transfer ratios α obtained on samples with various values of $CoSi_2$ base thickness, at room temperature and at 77 K. Error bars correspond to the dispersion of values for different devices fabricated on the same wafer. The results clearly show that Eq. (4) is verified, with values of λ_b of 8 ± 1.5 nm at 300 K and 35 ± 5 nm at 77 K. This agreement is in strong favor of ballistic theory. Moreover, the λ_B values are close to the mean free path deduced from resistivity measurements /22/. This indicates that the same scattering mechanisms control both the electron transport close to the Fermi level and the hot electron relaxation for energies in the 0.7 eV range above E.

Figure 11 shows α versus $V_{BC}^{-1/4}$ curves, taken at different temperatures, where V_{BC} is the base-collector bias. Equation 5 clearly holds since $V_{BC}^{-1/4}$ is directly proportional to x_m /21/. Furthermore, a mean free path is extracted from the slope of the curves and its temperature variation is shown in the inset of Figure 10. The very low values obtained for λ_{ph} (< 2 nm) remain to be understood.

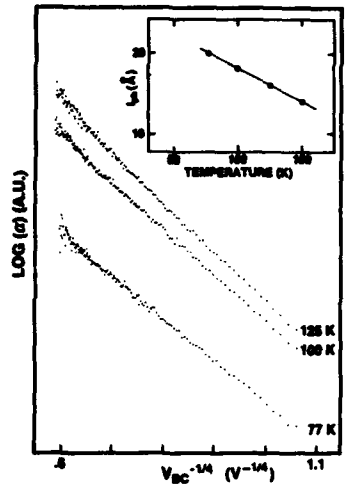


Fig. 11 Transfer ratio versus base-collector bias at different temperatures. The mean free path λ_{ph} deduced is shown in the inset as a function of temperature /22/.

Another problem requiring explanation is the overall low values of α_o (~ 0.3 at RT) corresponding to α/α_c ~ 0.6. The answer is most probably related to the quantum nature of the electron. Indeed, the electron energy E_1 in the metal is in the 5 eV range while its value E_2 in the semicondutor is in the 50 meV range, i.e. the Schottky barrier lowering. Consequently, the abrupt change in the electron wavelength leads to a quantum reflection at the metal-semiconductor interface. If the crude model of the abrupt-step barrier is assumed, the quantum transmission coefficient α_q is /22/

$$\alpha_q = 1 - [(m_2^* E_1)^{1/2} - (m_1^* E_2)^{1/2}/((m_2^* E_1)^{1/2} + (m_1^* E_2)^{1/2})]^2 \qquad (6)$$

where m_i^* is the effective mass in the i^{th} medium. Taking $m_1^* = 1$ in $CoSi_2$ and $m_2^* \sim 0.3$ in Si, one obtains $\alpha_q \sim 0.5$, which is in fair

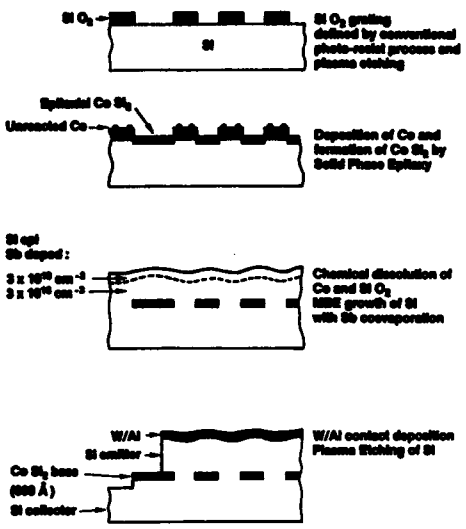


Fig. 12 Main fabrication sequence of epitaxial Si/$CoSi_2$/Si permeable base transistor /27/.

agreement with the experimental data of $\alpha_o \sim 0.3$ taking into account the collector backscattering.

All these results, as well as those obtained by Sze and his coworkers /21/, show that the quantum reflection is a severely limiting factor for the device interest of SMS transistors. These results suggest, in order to reduce this reflection, the use of highly assymmetrical SMS structure, for instance by use of two different semiconductors and/or metals.

V. Si/$CoSi_2$/Si Permeable Base Transistors

The idea of Permeable Base Transistor (PBT) is to take advantage of mechanism 2 described in the preceeding section. The main advantages of this device are :

- For small enough Si channels, PBT's behave like thermionic devices so that their transconductance is very high /23/.
- There are no fundamental limitations such as Quantum Reflection in SMS transistors.
- It is easily shrinkable with no problems of punchthrough such as in bipolar transistors.

Two ways are possible to imbed semiconductor channels in a metal grid.

1. The first one is to use the natural porosity of a metal layer on the surface of a semiconductor, leading to a Natural Permeable Base Transistor (also called Metal Grid Transistor /24/). Discontinuous $CoSi_2$ films /17, 10/ but also W layers deposited during Si CVD growth /24/ have been used. However, the high input capacitance as well as the lack of control in the geometry of metal openings are not in favour of such a device.
2. The second way is to define the opening by lithographic techniques. Bozler and Alley /25/ were the first to realize such a structure. They have used W grids of 320 nm periodicity and CVD deposited GaAs as the semiconductor, though the W/GaAs system is not epitactic.



Fig. 13 Backscattered electron image in SEM of epitaxial $CoSi_2$ lines, 0.3 µm wide (bright areas) buried in Si lattice.

Si/$CoSi_2$/Si Permeable Base Transistor with micronic Si channel dimensions have been independently obtained by Ishiwara et al /26/ and Rosencher et al /27/. Submicron Si/$CoSi_2$/Si PBT are currently under study in few laboratories. Though the problems of materials associated to epitaxial growth in submicron structures are far beyond the scope of this paper, we show in Figure 12 the main fabrication sequence of epitaxial Si/$CoSi_2$/Si permeable base transistor. Figure 13 shows the backscattered electron image in a Scanning Electron Microscope of a 300 nm $CoSi_2$ grid imbedded in a Si lattice. Electrical behaviour of such PBT's is currently under study.

VI. Conlusion

The physics and technology of metal-semiconductor heterostructures are clearly a rapidly expanding new field of research. Many of the experimental results described in this review paper are still in their early stage. A non limitative list of unresolved problems and future developpment is given hereafter :

- Mechanism of $CoSi_2$ growth on top of Si (role of the early stage nucleation, roughness, planarisation, influence of strains...).
- Microscopic nature of the $CoSi_2$/Si interface regarding to parallel transport properties.
- 2D localisation effects.
- Scattering mechanisms in SMS transistors (hot electron in the metal, Si collector backscattering...).
- Other epitaxial metals on Si such as ternary silicides.

References

/1/ **S. Saitoh, H. Ishiwara, S. Furukawa** : Appl. Phys. Lett. 37, 203 (1980).

/2/ **J.C. Bean, J.M. Poate** : Appl. Phys. Lett. 37, 643 (1980).

/3/ **A. Ishizaka, K. Nakagawa, Y. Shiraki** : In Proceedings of the Symposium on molecular beam epitaxy and clean surface techniques, (The Japan Society of Appl.Phys., Tokyo, 1982), p. 183.

/4/ **F. Arnaud d'Avitaya, S. Delage, E. Rosencher, J. Derrien** : J. Vac. Sci. Technol. B3, 770 (1985).

/5/ **E. Rosencher, S. Delage, F. Arnaud d'Avitaya** : J. Vac. Sci. Technol. B3, 762 (1985).

/6/ **K. Ishibashi, S. Furukawa** : Appl. Phys. Lett. 43, 660 (1983).

/7/ **B.D. Hunt, N. Lewis, E.L. Hall, L.G. Turner, L.S. Schowalter, M. Okamoto, S. Hashimoto** : In Proceedings of the MRS Conference, 37, "Layer Structures, epitaxy and Interfaces", Ed. J.M. Gibson and L.R. Dawson, p. 131.

/8/ **C. d'Anterroches, F. Arnaud d'Avitaya** : Thin Solid Films 137, 351 (1986).

/9/ **E. Rosencher, P.A. Badoz, C. d'Anterroches, G. Glastre, G. Vincent, F. Arnaud d'Avitaya** (to be published).

/10/ **R.T. Tung, A.F.J. Levi, J.M. Gibson** : Appl. Phys. Lett. 48, 635 (1986).

/11/ **F. Arnaud d'Avitaya, J.A. Chroboczek, G. Glastre, Y. Campidelli, E. Rosencher** : J. of Crystal Growth (to be published).

/12/ **J.C. Hensel, R.T. Tung, J.M. Poate, F.C. Unterwald** : Phys. Rev. Lett. 54, 1840 (1985).

/13/ **P.A. Badoz, A. Briggs, E. Rosencher, F. Arnaud d'Avitaya** (to be published).

/14/ **R. Madar, A. Briggs** (to be published).

/15/ **E. Rosencher, P.A. Badoz, A. Briggs, Y. Campidelli, F. Arnaud d'Avitaya** : In Proceedings of the first international symposium of silicon molecular beam epitaxy, ed. J.C. Bean, The Electrochemical Society, Vol. 85-7, p. 268 (1985).

/16/ **E. Rosencher, S. Delage, Y. Campidelli, F. Arnaud d'Avitaya** : Electron. Letters 20, 762 (1984).

/17/ **E. Rosencher, S. Delage, F. Arnaud d'Avitaya, C. d'Anterroches, K. Belhaddad, J.C. Pfister** : Physica B134, 106 (1985).

/18/ **J.C. Hensel, A.F. Levi, R.T. Tung, J.M. Gibson** : Appl. Phys. Lett. 47, 151 (1985). See also Ref. 10.

/19/ **E. Rosencher, P.A. Badoz, J.C. Pfister, F. Arnaud d'Avitaya, G. Vincent, S. Delage** : Appl. Phys. Lett. 49, 271 (1986).

/20/ **J.C. Pfister, E. Rosencher, K. Belhaddad, A. Poncet** : Solid State Electron. 29, 907 (1986).

/21/ **S.M. Sze** : In Physics of Semiconductor Devices, (Wiley-Interscience, New York, 1969) Chap. 11.

/22/ **P.A. Badoz, E. Rosencher, S. Delage, G. Vincent, F. Arnaud d'Avitaya** : In Proceedings of the 18th International Conference of Physics of Semiconductors (1986, Stockholm) (to be published).

/23/ **A. Marty, J. Clarac, J.P. Bailbe, G. Rey** : IEE Proc. 130, 24 (1983).

/24/ **J. Lindmayer** : Proc. IEEE, 1751 (1964).

/25/ **C.O. Bozler, G.D. Alley** : IEEE Trans. Electron Devices 27, 1128 (1980).

/26/ **K.Ishibashi, S. Furukawa** : IEEE Trans. on Electron Devices 33, 322 (1986).

/27/ **E. Rosencher, G. Glastre, G. Vincent, A. Vareille, F. Arnaud d'Avitaya** Electron. Letters 22, 699 (1986).

Session A4.2

Integrated Optoelectronics II

Chairman: V. Ghergia

Thursday, September 17, 1987

PROCESS AND MODEL OF SHORT - GATE DIFFUSED InGaAs JFET's FOR INTEGRATED PIN - FET PHOTODETECTOR

L.Nguyen, M.Allovon, P.Blanconnier, B.Bourdon *
E.Caquot, A.Scavennec

Centre National d'Etudes des Telecommunications
Laboratoire de Bagneux, 196, avenue Henri Ravera
92220 Bagneux France
* *Laboratoires de Marcoussis CRCGE*
Rte Nozay 91460 Marcoussis France

ABSTRACT : Junction Field Effect Transistors with a 1 micron diffused gate have been fabricated on InGaAs/InP grown by molecular beam epitaxy (MBE). Transconductances higher than 200 mS/mm with a channel doping level of 2.10^{16} cm^{-3} have been measured. A compact model for PIN - FET photodetector has been developed to optimize the signal to noise ratio.

INTRODUCTION

For 1.55 μm optical link photodetection, integrated PIN - FETs appear to be a very attractive solution in order to improve the signal to noise ratio and detection threshold. Integrated with an InGaAs photodiode, InGaAs JFET is a technological response for this need. Among the various available devices [1], diffused or implanted JFETs are easy to process and retain interest in terms of insensitivity to surface - related phenomena (surface depletion, surface stability) and in terms of gate leakage current.

The aim of this paper is to report on the process and results on such a short - gate diffused JFET as it is now developed at CNET. The associated model will be emphasized in order to enable the design of an integrated PIN - FET photodetection circuit.

1. InGaAs Zn - DIFFUSED GATE J - FET TECHNOLOGY :

Starting from a InGaAs layer grown by MBE (Molecular Beam Epitaxy) on a semi - insulated substrate, the JFET's process technology gives the structure shown on figure 1. The transistor active region is delimited by mesa, the gate is a P - N junction obtained by a localized diffusion through a SiN mask. The N - type and P - type metalizations are successively aligned with the diffused zone for the source - drain and gate contacts.

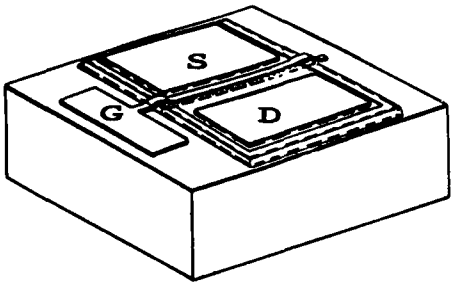


FIGURE 1

Structure of the diffused JFET

1.1. InGaAs layer grown by MBE

InGaAs layers are grown in a Riber 2300 equipment with a rotating substrate holder. Epitaxial growth is conducted at 500 C, with a growth rate of 1 μm/hr. The cristallographic properties are usually good with a lattice mismatch less than $\pm 5.10^{-4}$. At a

doping level of 2. 10^{16} cm^{-3}, mobilities of 9000 and 24000 cm^2/Vs are measured at 300 and 77 K respectively. A sharp decrease of the electron mobility and doping level in the InGaAslayer is sometimes observed close to the interface with S.I. substrate. To improve the transistor channel properties, an intrinsic quaternary layer (GaAlInAs) has been grown as buffer on several slices. The FET InGaAs layer is typically doped (Si) at 2.10^{16} cm^{-3}, with a thickness of 1.2 microns.

1.2. J - FET fabrication

Mesas are formed in the InGaAs layer by chemical wet etching (H3PO4:1 H2O2:1 H2O:8), with an etching rate of 5000 Angstrœms/ mn. The diffusion mask is deposited : 2000 Angstrœms of SiN by PE - CVD on both InGaAs mesa and InP substrate. Mask opening are realized by plasma etching (CF4). Diffusion is conducted in a sealed ampoule, evacuated to 10^{-7} torr, at 500°C. Zn is provided by ZnAs2. The diffusion depth is 0.8 micron. Deposited by electron beam evaporation, ohmic contacts are delineated by the lift - off technique. AuGeNi/Ag/Au contacts for source and drain are alloyed by halogen lamp (430°C for 10 s). TiAu contact is deposited for the metalization gate. The narrowest dimensions are 1 micron for the SiN opening, 3 microns for the metalization gate between 5 microns drain - source spacing.

2. J - FET CHARACTERISTICS

The best results in terms of transconductance have been obtained on enhancement mode devices from a slice with a doping level close to the interface lower than anticipated. Average transconductance of 175, 138, 67, 42 mS/mm have been measured for gate lengths of 1, 1.5, 2.5, 4 microns respectively at a gate voltage of 0.7 V and for a drain to source voltage of 1.4 V. DC characteristics of a device with a 1x300 microns gate are shown on figure 2. Measurements of the gate capacitance at 0 V give 0.85 pF for a 1.5x300 micron² transistor. Source and drain resistances are quite low : 0.2 Ω/mm gate width, with a contribution of the contact resistance of 0.1 Ω/ mm (resistivity contact values 10^{-7} $\Omega.cm^2$). Gate resistances are 1 Ω/mm, with a contact resistivity of 10^{-5} Ωcm^2. Usually gate junction characteristics show a negligeable leakage current, about 100 nA at - 3 V.

3. J - FET MODELLING

Modelling is required for the PIN - FET structure and design optimization in order to obtain a high signal/noise ratio. The model can be also used in the parametrical test of the real devices. An analytical approach has been chosen to permit an easy implementation on desk computers. The model is based on the classical two - regions model [2], with specific parasitic effects of the geometry of the gate (diffused or etched P - region).

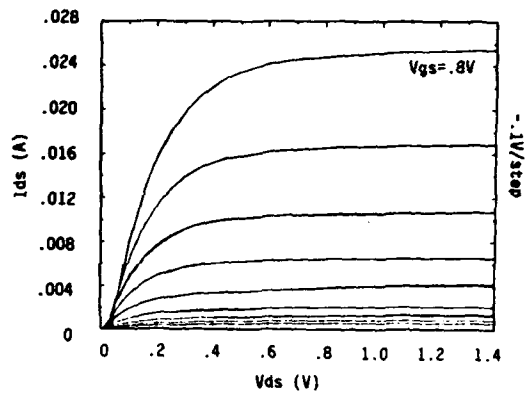


FIGURE 2

Experimental Id (Vds) characteristics of a 1 μm gate length JFET. W = 200 microns, Nd = 2.10^{16} cm^{-3}.

3.1. Cylindrical junction

The diffused region section can be represented by a rectangle where the length is the diffusion mask opening length (1 micron), the width is the diffusion depth (0.8 micron), and with two rounded parts at the edges due to the lateral diffusion under the mask (figure 3). These quasi circular regions induce specific depletion regions whose radius can be computed from the expression of the electric field E(r) of a cylindrical junction of radius r :

$$E(r) = \frac{1}{\varepsilon_0 \varepsilon_r r} \int_{r_j}^{r} r\,\rho(r)\,dr + \frac{\text{constant}}{r}$$

where $\rho(r)$ is the electronic charge at radius = r

Assuming Na, P - type doping level higher than Nd and with the appropriate boundary conditions, we can deduce the potential in the circular depletion region :

$$V_{cyl}(r) = \frac{q N_D r_j^2}{2 \varepsilon_0 \varepsilon_r} \left(1 + \frac{r}{r_j}\right)^2 \left(\ln\left[\left(1 + \frac{r}{r_0}\right) - 0.5\right] + 0.5\right)$$

The lateral extension rdep of this depletion region at drain side is controlled by the potential Vgd. If Vb is the built - in voltage of the junction, rdep is given by the equation :

$$V_{cyl}(r_{dep}) = V_B - V_{GD}$$

The same procedure is used to compute the lateral extension of the depletion region at source side.

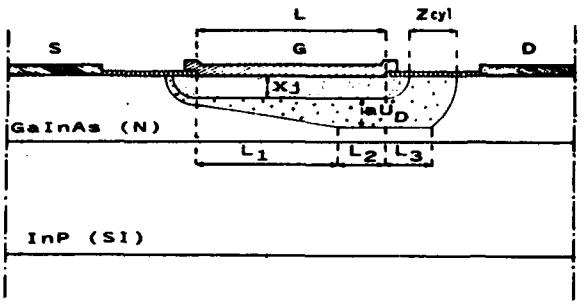


FIGURE 3

Cross – section of the 2 – D depletion region of the JFET. Lateral diffusion is responsible of cylindrical depletion regions at the edges of the diffused gate.

3.2. Channel regions

As for the basic two – region model, the channel is shared in two parts (see figure 3) : the first one (L1) where the mobility is constant, and the second one where the velocity is constant. This second part is shared in two parts also : the length L3 is due to the influence of the cylindrical depletion region and L2 is the length of the saturated channel due to the vertical depletion region. If zcyl is higher than the vertical depletion region height (a ud), the length L3 is given by :

$$L_3 = \sqrt{(x_j + z_{cyl})^2 - (x_j + au_D)^2}$$

Otherwise L3 = 0

3.3. Current calculation

The current Ids is given by the following equation :

$$Ids = \frac{Vp}{Ro}\frac{L}{L_1}\left(3\,(u_D^2 - u_S^2) - 2\,(u_D^3 - u_S^3) \right)$$

where Vp is the pinch – off voltage, Ro the resistance of the channel, L the channel length, and u_S and u_D are given by :

$$u_D = 1 - \frac{IdsRo}{E_C L_1}$$

$$u_S = \left(\frac{-Vgs + Vb + RsIds}{Vp} \right)^{1/2}$$

Further analysis leads to :

$$V_{chan} = Vp\,(u_D^2 - u_S^2) \frac{2E_C a u_D}{\pi} \sinh\left(\frac{\pi(L_2 + L_3)}{2au_D}\right)$$

where Vchan is the voltage drop along the channel, a the channel thickness.

3.4. Exploitation of the model

The model can calculate the two main elements for the signal to noise ratio Cgs: the gate – source capacitance and gm: the transconductance. The S/N ratio at the input can be written as :

$$S/N = \frac{(S_0 P)^2}{\int \Delta F \left(2\,q[i_0 + S_0 P] + 4\,kT/R + 2\,kT\,\Gamma\,C^2\,w^2/g_m\right) df}$$

where S is the photodiode sensitivity and P the optical power. $i_o$is the total input leakage current,R the biasing resistor, the channel noise factor and C the total input capacitance (Cdiode + Cfet). For the large bandwidth envisioned (ΔF 1 GHz), the signal/noise ratio is mostly governed by the factor C^2/gm. Figure 4 shows for instance the variations of the minimum detectable power with the channel doping level.

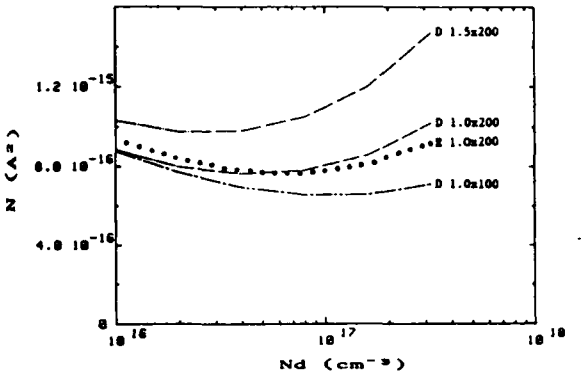


FIGURE 4

Variation of the photodetector input noise, assuming a capacitance of 0.1 pF for the phototdiode. Parameters : electron mobility (at 10^{16} cm^{-3}) = 8000 cm^2/Vs, saturation velocity = 2.5 10^7 cm/ s, Vds = 4 V. D : diffused gate, E : etched gate. 1x200 : gate dimensions (in microns).

4. CONCLUSION

JFET with gate junction made by Zn diffusion have been successfully realized on MBE grown InGaAs layers. Values larger than 200 mS/mm for the transconductance have been measured on 1 micron gate length, enhancement mode devices. The processing technology is compatible for the integration with a photodiode. An analytical model has been used to determine the influence of the transistor structure, and its parameters such as channel doping level and gate dimensions, on the PIN – FET signal to noise ratio.

REFERENCES

[1] T.Y. Chang, et al, *"Junction field effect transistor using InGaAs material grown by Molecular Beam Epitaxy."*, IEEE El. Dev. Lett., EDL 3 (1982), 56.

[2] T. Grebene, S.K. Gandhi, *"General theory for pinched operation of the junction – gate FET's."*, Solid – State Electron., 12, (1969), 573.

MONOLITHIC INTEGRATION OF A GaInAs JFET AND A GaInAs PHOTODIODE

P.J.G. DAWE, D.A.H. SPEAR, G.H.B. THOMPSON and G.R. ANTELL

STC Technology Ltd.
London Road, Harlow, Essex, CM17 9NA, England

Monolithically integrated GaInAs/InP pin/JFETs have been made from LPE grown material. The method of integration allows the pin and JFET to be optimised independently, while keeping stray capacitance extremely low. Similar JFETs have also been made by MOVPE.

1. INTRODUCTION

Monolithic integration of photodetectors and FETs in optical receivers is advantageous for improving the high frequency sensitivity by reducing stray capacitance [1] - see Figure 1, for simplifying wavelength multiplex systems, for obtaining better reproducibility and eventually for opening the way to low cost, high performance wideband multi-channel optical communication systems. For the low fibre loss optical windows at 1.3 and 1.5 μm wavelengths, the conventional detector material of epitaxial GaInAs on an InP substrate can also be developed to provide a good basis for

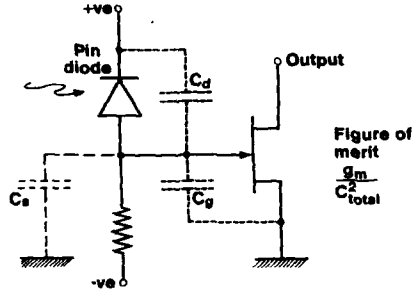


FIGURE 1

pin/FET receiver circuit

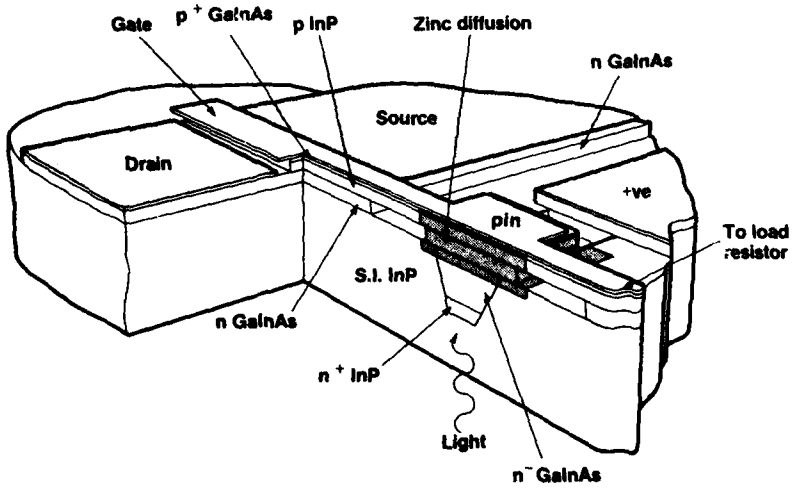


FIGURE 2

Drawing of integrated pin/FET

junction FETs. To integrate the two types of devices successfully on the same substrate it is necessary to find a means of reconciling their very different requirements in terms of the thickness and doping level of the epitaxial layers, and to adapt the diode to a semi-insulating substrate. We have achieved integration by fabricating the photodiode in material embedded by LPE growth in SI InP [2] and combining it with a simply processed JFET which uses additional LPE layers of n-type GaInAs for the channel and p-type InP for the gate. The drawing of Figure 2 shows one convenient way of blending the p-side of the pin diode with the p-InP that forms an extension of the gate of the JFET, using Zn diffusion.

2. GROWTH

This structure has been realised by a two-stage LPE process. The first stage is the growth of the n^+ InP back contact layer and the n^- GaInAs detector layer, and an etch stage to give bars of detector material embedded in the semi-insulating substrate. The second stage is the growth of a four-layer JFET structure:

Contact layer:

0.1 μm GaInAs, $p = 2.10^{18} cm^{-3}$

Gate layer:

0.3 μm InP, $p = 6.10^{17} cm^{-3}$

Anti-melt-back layer:

0.05 μm GaInAsP, $\lambda = 1.54$, undoped

Channel layer:

0.2 μm GaInAs, $n \cdot 6.10^{16} cm^{-3}$

Figure 3 shows the good planarity of this process.

3. PROCESSING

A masked zinc diffusion forms the p-electrode of the photodiode. The gate pattern (Ti/Pt/Au) is defined by the lift-off process. The gate is etched using selective wet etches which under cut the metal. The source and drain contacts (AuGe/Ni/Au) are evaporated to form self-aligned (Figure 4) and non-self-aligned JFETs. A further wet etching stage, using a photoresist mask, is used to isolate the devices. A completed pin/FET is shown in Figure 6.

4. PERFORMANCE

The measurements in this paragraph were made on individual (not integrated) devices. We find a photodiode capacitance down to 60 fF, a dark current of 35 nA and quantum efficiency greater than 50% in the range 1.3-1.5 μm [1]. For the JFETs we find a transconductance up to 95 mS/mm (Figure 5), and a junction capacitance of 1.7 pF/mm (measured at $V_g = 0$ on a device with gate length about 2.5 μm). This gives a creditable ratio C/g_m of 18 ps - this is a more valid measure of usable performance than a transconductance measurement. We estimate that the internal transit time is about 15 ps, the channel mobility is over 6000 $cm^2/v.s$ and the peak velocity about 2.5×10^7 cm/s. These figures show that GaInAs is an excellent material for high speed transistors.

The integrated pin/FETs processed have had rather lower transconductance and higher photodiode capacitance than the discrete examples measured, because of non-optimum layer thickness and doping levels. However, the integration scheme appears not to impose a penalty on the performance of the components.

Integrated pin/FETs have been mounted on ceramic and back illuminated using chopped light from a monochromator. A circuit as in Figure 1 was used. A voltage gain of 5 to 8 was obtained in transimpedance operation. No high frequency measurements have yet been made, although we estimate that the JFET cut-off frequency is over 6 GHz.



FIGURE 3

Four-layer FET growth on embedded pin structure

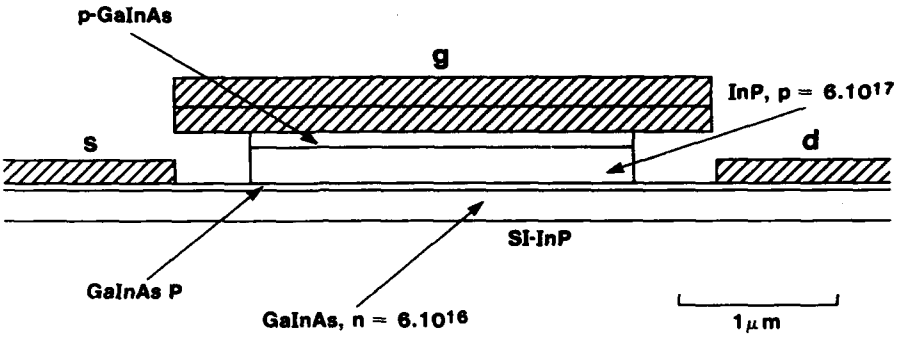


FIGURE 4

Cross-section of self-aligned JFET

5. MOVPE JFET

We have also realised the same JFET design (without the anti-melt-back layer) by MOVPE and the results are very similar. This work is at an early stage and such JFETs have not been integrated with photodiodes.

6. CONCLUSION

Pin photodiodes and GaInAs/InP JFETs, both specifically designed for integration, have been made by an LPE-based process and measured, and show good performance. Integrated pin/FETs have been made and tested. It appears that the integration scheme does not degrade the performance of the individual components.

MOVPE-grown JFETs show similar performance to the LPE-grown devices.

ACKNOWLEDGEMENT

This work is supported by the EEC through ESPRIT Project 263.

REFERENCES

[1] Smith, R.G. and Personick, S.D., Receiver design for optical fiber communication systems in: Kressel, A. (ed.) Semiconductor Devices for Optical Communication (Springer-Verlag, Berlin, 1982).

[2] Dawe, P.J.G., Spear, D.A.H. and Thompson, G.H.B., Planar embedded GaInAs photodiode on semi-insulating InP substrate for monolithic integration, Elect. Lett. 22, pp. 722-724, 1986.

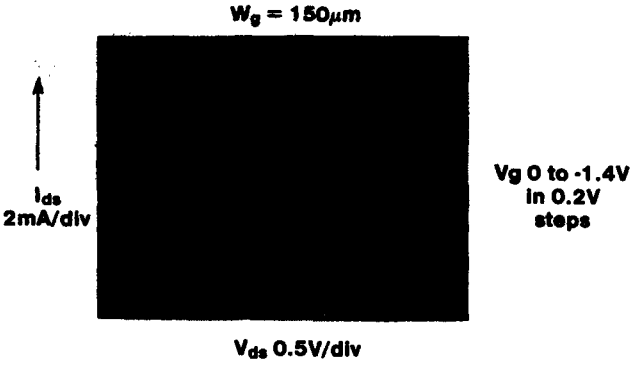


FIGURE 5

Transfer characteristic of self-aligned JFET

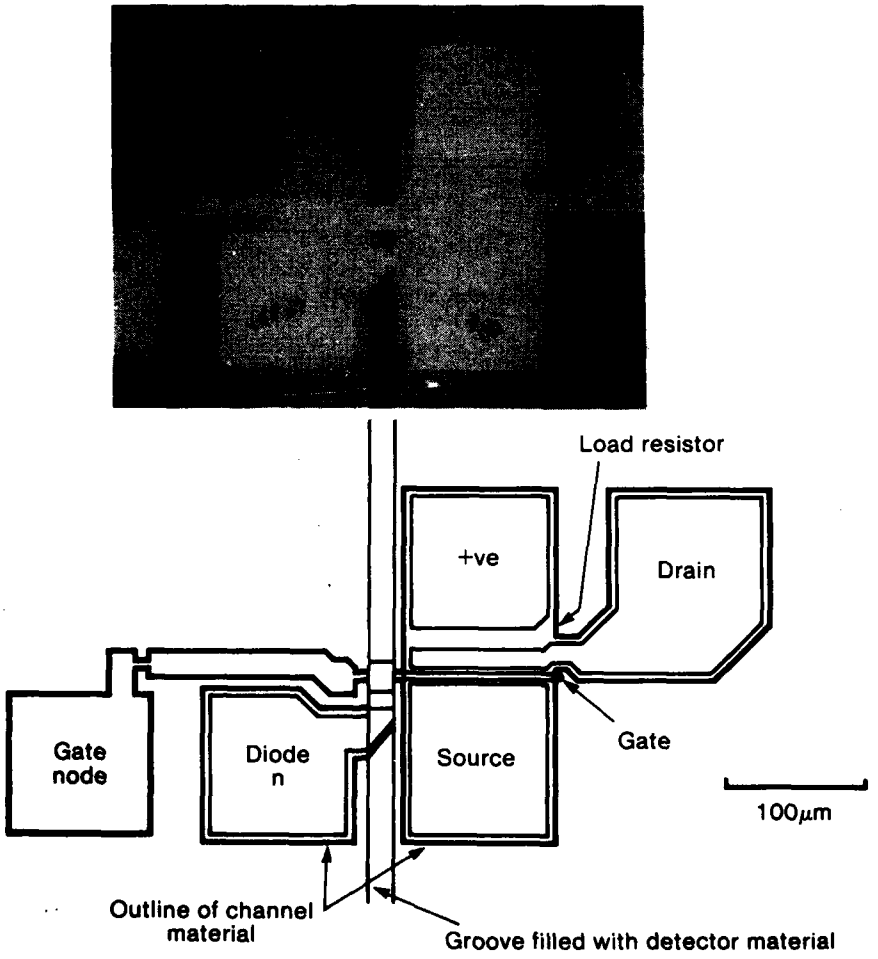


FIGURE 6

A completed integrated pin/FET

InGaAs/InP SAM Photodetectors Fabricated by Structure Controlled Acceptor Diffusion

E.Kühn, N.Klötzer, H.W.Marten, A.Schmiech

4. Physikalisches Institut der Universität Stuttgart
D - 7000 Stuttgart 80, Federal Republic of Germany

The Cd-diffusion at InP/InGaAs heterojunctions was investigated by C-V profiling of p-n junctions located on the InP-side close to the heterointerface. A compensated layer in front of the heterojunction is observed. Reducing the distance p-n junction-heterointerface decreases the width of this layer causing a distinct drop of the breakdown voltage of the junction. This effect is used to realize SAM APD's with guard ring by a structure controlled modulation of the distance p-n junction-heterointerface. Avalanche gain up to M = 15 is achieved. The breakdown voltage in the multiplication area is at least 40% smaller than in the guard ring area. A background doping level N_D=2-3·10^{16}cm^{-3} in the InP-layer is required for optimum detector operation.

1) INTRODUCTION

Designing efficient InP/InGaAs SAM avalanche photodetectors for long wavelength optical fibre communication (1.3-1.6μm) [1] requires a well defined optimization of the detector stucture. An important topic of design is the electric field distribution within the absorption and the multiplication layer. In order to ensure sufficient avalanche gain the minimum electric field strength E = 4.5·10^5V/cm has to be achieved at breakdown voltage in the InP layer [2,3]. However, in order to keep tunneling currents from the ternary layer negligible the electric field must not exceed 1.9·10^5V/cm at the heterointerface [4,5].

Surface leackage currents and surface breakdowns can be avoided by a so-called guard ring which is mostly realized by a double diffusion process [6], or diffusion through a SiO_2 barrier [7], or the so-called self-guard ring effect [8].

However, in the case of p-n junctions prepared by acceptor diffusion close to the heterointerface there is a lack of detailed knowledge concerning the exact impurity profile. Thus the design of efficient detector structures is rather difficult. In this work we report on the influence of InP/InGaAs heterojunctions on Cd-diffusion profiles. Based on these studies we succeeded in preparing guard ring APD's by a *single*, structure controlled diffusion process.

2. Cd-DIFFUSION PROFILES

For homogeneous III-V compounds such as InP with background doping levels N_D < 5·10^{16}cm^{-3} the acceptor diffusion leads to p^+-p-n structures [9]. The impurity profiles reveal two diffusion fronts i.e. a steep decay of the acceptor concentration at the p^+-p junction with an adjacent slightly graded (error function shaped) impurity tail.

In the case of p^+-p-n junctions located close to abrupt InP/InGaAs heterojunctions drastically different impurity profiles are observed which is briefly outlined below (for a detailed discussion see [10]). Fig. 1a shows the apparent carrier concentration profiles of p-n junctions with the p^+-p junction located 3.5μm, 2.2μm, 1.8μm, and 1.5μm in front of a InP/InGaAs heterojunction. The profiles (fig. 1b) were obtained by C-V profiling of mesa diodes prepared from a Cd-diffused InP/InGaAs/InP heterostructure with a bevelled InP top layer. Due to the steep slope of the impurity concentration at the p^+-p junction these profiles actually represent the net carrier concentration on the n-side of the p-n junction.

The profiles reveal a low net carrier concen-

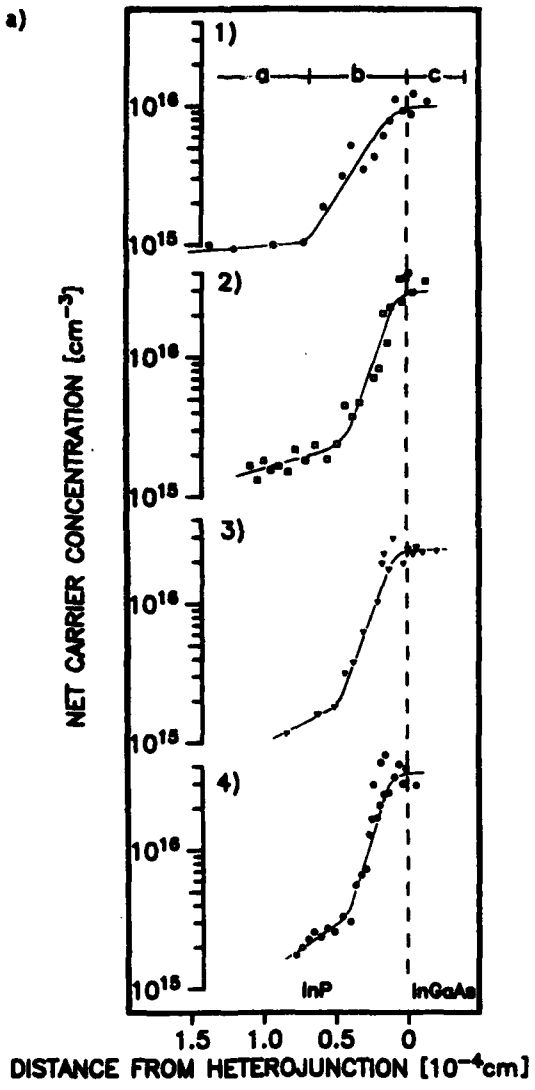


Figure 1

a) Net carrier density profiles of mesa diodes prepared by Cd-diffusion into an InP/InGaAs/InP sample with bevelled InP top layer. Profile 1-4: (mean) distance p^+p-junction - heterojunction decreasing from 3.5μm to 1.5μm. InP background doping level N_D=7·10^{15}cm^{-3}, T_{diff}=610°C, t_{diff}= 30min, bevelling angle φ=0.25°.
b) Schematic cross section of the sample.

tration compensated to approximately 1/5 of the background doping level (profile 1, region *a*) followed by a steep increase nearly up to the background doping level (region *b*) which is caused by a decay of the Cd-concentration in front of the heterojunction. The position of the heterointerface is determined by comparing these n(x) profiles with non-bevelled mesa diodes where the accumulation and depletion layer of the heterojunction can be clearly identified. Within the ternary layer (region *c*) the net carrier density is equal to the background doping level. The compensation factor mentioned above depends on the background doping level and decreases to 1/2 for $N_D = 2 \cdot 10^{16}$cm^{-3}. With decreasing distance p^+-p junction - heterojunction (profiles 2, 3, and 4) the width of the compensated region *a* decreases from 3.0μm to 1.3μm whereas the net carrier concentration gradient increases from $5 \cdot 10^{18}$cm^{-4} to $8 \cdot 10^{19}$cm^{-4}. Thus a reduction of the breakdown voltage has to be expected in analogy to a p^+-ν-n junction with decreasing ν-layer width. This correlation between Cd-impurity profiles and breakdown voltages might also explain the 'self-guard ring' effect reported in [8].

3. SAMPLE PREPARATION

SAM avalanche photodiodes with guard ring were fabricated using InP/InGaAs double heterostructures: S-doped substrate, 4.5μm LPE InGaAs layer ($N_D = 2 \cdot 10^{16}$cm^{-3}), 5.1μm VPE InP layer ($N_D = 1.5 \cdot 10^{16}$cm^{-3}). In order to prepare guard ring and multiplication area simultaneously by a single diffusion process, cylindric cavities (diameter 60μm, depth 1.5μm) were etched into the InP top layer (H_2SO_4:H_2O_2:H_2O = 8:1:1, 60°C). Afterwards different diffusion experiments were carried out at T_{diff} = 610°C using a Cd_3P_2 + P_4 source. The diffusion time was varied in the range 40min ≤ t_{diff} ≤ 60min. Due to the surface structure the diffusion front is located closer to the heterointerface below the cavities. Thus the breakdown voltage V_b of this (multiplication) area is reduced compared to V_b of the surrounding (guard ring) area. Since the breakdown voltage and therefore the multiplication properties are ex-

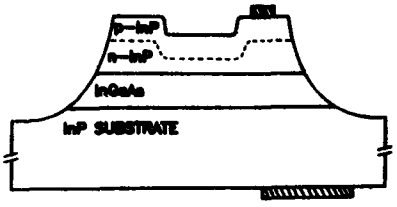


Figure 2

Schematic cross section of a guard ring SAM APD. Dashed line: p-n junction. The reduction of the distance p-n junction - heterointerface below the cavity results in a decreased breakdown voltage (multiplication area).

tremely sensitive to slight variations of the distance p-n junction - heterointerface, the uniformity of the avalanche gain essentially depends on the flatness of the cavity bottom: with the above given H_2SO_4-etchant the uneveness is restricted to the range $\Delta < 70nm$. After evaporation of ohmic contacts APD's were prepared by etching concentric mesas around each cavity (diameter 100μm, $HBr:H_2O_2:H_2O$ = 10:1:10, see fig. 2).

4. DEVICE PROPERTIES

Avalanche gain up to M = 15 is achieved (t_{diff} = 40min). With increasing diffusion time the maximum multiplication factor decreases to M = 3.5 (t_{diff} = 60min) presumably caused by a higher voltage drop across internal series resistances due to increasing dark currents [11].

The absence of local breakdowns (microplasma effects) was proved by spatially resolved photocurrent measurements as well as noise measurements in the 100kHz - 1MHz frequency range revealing white photocurrent noise and multiplication noise spectra.

The high frequency response was measured using a Nd:YAG laser (λ = 1.06μm, 120ps pulse length). Pulse rise times $\tau \leq$ 125ps and half widths $\Delta\tau \leq$ 165ps are achieved at M = 4 (t_{diff} = 50min).

Typical photocurrent-voltage characteristics for the three different diffusion series (t_{diff} = 40min, 50min, and 60min) are shown in fig. 3. As predicted in chapter 2 the onset of avalanche multiplication is reduced to lower voltages with decreasing distance p-n junction - heterointerface. Simultaneously the breakdown voltage decreases from 57V (t_{diff} = 40min) to 37V (t_{diff} = 60min). The dependence of the breakdown voltage on the distance Δd between the p-n junction and the heterointerface was found to be very sensitive for low values of Δd (fig. 4).

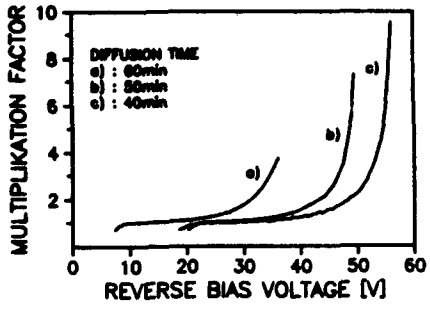


Figure 3

Photocurrent multiplication versus reverse bias for different distances p^+-p junction - heterojunction (t_{diff}=40min, 50min, and 60min).

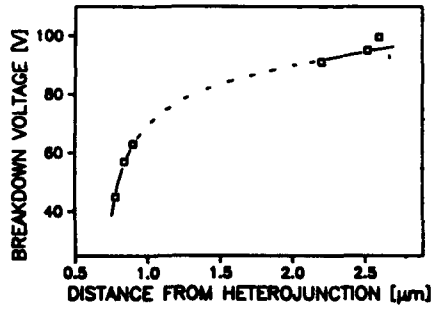


Figure 4

Dependence of the breakdown voltage on the distance p^+-p junction - heterointerface.

For all samples the breakdown voltage of the multiplication area proved at least 40% smaller compared to the guard ring area. Further enhancement of the maximum avalanche gain without loss of the guard ring effect should be achieved by increasing the distance p-n junction - heterointerface in the multiplication area.

5. DISCUSSION

As mentioned above sufficient avalanche gain

requires $E \geq 4.5 \cdot 10^5$V/cm at the p-n junction, i.e. the charge density $\sigma = 3 \cdot 10^{12}$cm^{-2} has to be depleted from both the InP *and* the InGaAs portion of the space charge layer. In order to avoid significant tunneling currents the charge density depleted from the InGaAs layer must not exceed $\sigma = 1 \cdot 10^{12}$cm^{-2} [12]. Fig. 5 shows the apparent

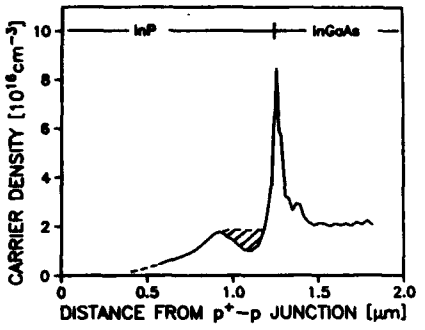


Figure 5

Carrier density profile of the multiplication layer of a SAM APD as shown in fig. 2 (t_{diff}= 40min, guard ring removed). Dashed line: InP-background doping level. Hatched area: intrinsic space charge of the heterojunction.

carrier density profile of a mesa diode (t_{diff} = 40min, guard ring area removed). The position of the heterointerface is determined from the depletion and accumulation layer of the heterojunction. The carrier density swapped out of the space charge layer is obtained by integrating the n(x) profile on the InP side. Adding the intrinsic space charge of the depletion region of the heterojunction (fig. 5, hatched area) yields the total charge on the InP side. The result for our samples is $\sigma = 1.5 \cdot 10^{12}$cm^{-2} ($t_{diff}$ = 40min). Since tunneling currents become significant for $\sigma \leq 2 \cdot 10^{12}$cm^{-2} further optimization of the avalanche gain should be achieved by increasing the InP-background doping level to $N_D = 2\text{-}3 \cdot 10^{16}$cm^{-3}.

6. SUMMARY

Cd-diffusion into the InP-layer of InP/InGaAs heterostructures leads to a compensated layer in front of the heterojunction. Reducing the distance p^+-p junction - heterointerface decreases the width of this layer and simultaneously reduces the breakdown voltage. SAM avalanche photodetectors with guard ring were realized utilizing this effect. Due to the compensated layer an InP background doping level $N_D = 2\text{-}3 \cdot 10^{16}$cm^{-3} is required to avoid tunneling currents from the ternary layer.

REFERENCES

[1] Susa, N., Nakagome, H., Mikami, O., Ando, A., Kanbe, H., IEEE Journ. of Quant. Electron., QE-16 No. 8 (1980) 864

[2] Armiento, C.A., Groves, S.H., Hurwitz, C.E. Appl. Phys. Lett., 35 No. 4 (1979) 333

[3] Umebu, I., Choudhury, A.N.M.M., Robson, P.N. Appl. Phys. Lett., 36 No. 4 (1980) 302

[4] Ando, H., Kanbe, H., Ito, M., Kaneda, T. Jap. Journ. Appl. Phys., 19 No. 6 (1980) L277

[5] Cook, L.W., Tabatabaie, N., Tashima, M.M., Windhorn, T.W., Bulman, G.E., Stillman, G.E. Proc. Int. Symp. on GaAs and Related Compounds Inst. Phys. Conf. Ser. No. 56 (1980) pp. 361

[6] Ando, H., Susa, N., Kanbe, H. IEEE Transact. on Electron. Devices, ED-29 No. 9 (1982) 1408

[7] Ikeda, M., Wakita, K., Hata, S., Kondo, S., Kanbe, H., Electron. Lett., 19 No. 2 (1983) 61

[8] Taguchi, K., Matsumoto, Y., Nishida, K. Electron. Lett., 15 No. 15 (1979) 453

[9] Chin, A.K., Dutt, B.V., Temkin, H., Bonner, W.A., Roccasecca, D.D. Appl. Phys. Lett., 36 No. 11 (1980) 924

[10] Kühn, E., Marten, H.W., Schmiech, A. to be published

[11] Melchior, H., Lynch, W.T. IEEE Transact. on Electron. Devices, ED-13 No. 12 (1966) 829

[12] Kim, O.K., Forrest, S.R., Bonner, W.A., Smith, R.G., Appl. Phys. Lett., 39 No. 5 (1981) 402

MONOLITHIC INTEGRATION OF A $Ga_{0.47}$ $In_{0.53}$ As PHOTOCONDUCTOR WITH A n^-/n^+ GaAs RIB WAVEGUIDE : A SIMPLE DESIGN DEVICE

F.MALLECOT, J.P.VILCOT, D.DECOSTER, M.RAZEGHI*

Centre Hyperfréquences et Semiconducteur,
UA CNRS 287, Université des Sciences et Techniques de Lille Flandres-Artois
59655 Villeneuve d'Ascq Cedex, France.
*Laboratoire Central de Recherches, Thomson-CSF, Domaine de Corbeville
91401 Orsay Cedex, France.

We report the first fabrication of an optoelectronic integrated circuit constituted of a $Ga_{0.47}$ $In_{0.53}$ As planar photoconductive detector (suitable for 1.3µm - 1.55µm wavelength optical communication systems) associated with a n^-/n^+ GaAs rib waveguide. The capabilities of such a device to detect a part of a light propagating inside the waveguide are experimentally demonstrated.

Photodetectors directly integrated with optical waveguides are basic devices for optoelectronic integrated circuits (O.E.I.C.'s). Several structures have been previously proposed in Si, GaAlAs/GaAs, GaInAsP/InP [1] and more recently in GaInAs/InP [2], GaInAlAs/InP [3].

These last devices can be used for 1.3-1.5µm wavelengths signal detection. Nevertheless, it is well known that problems occur for the fabrication of field effect transistors on InP substrates. As a consequence, we present an other original solution which consists of the fabrication of a $Ga_{0.47}$ $In_{0.53}$ As photoconductive detector integrated with a GaAs rib waveguide using a GaAs substrate. The rib waveguide is made on a classical n^-/n^+ GaAs homojunction.

A schematic view of the device is given in figure 1. The optical waveguide is constituted of an undoped GaAs layer grown by M.B.E. on a n type highly doped GaAs substrate; the rib waveguide is obtained by an ion milling of 1.2µm of the undoped epilayer. The dimensions of the rib (6µm x 1.2µm) and the undoped epilayer thickness (4µm) have been calculated in order to provide a monomode propagation for 1.3-1.55µm wavelength optical signals. A

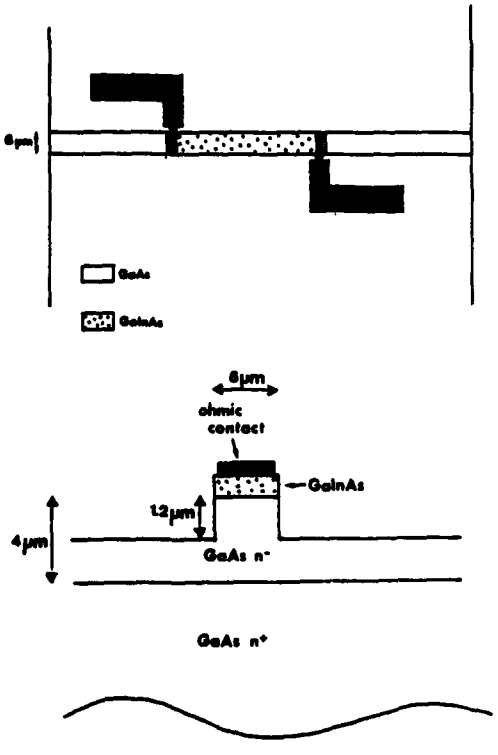


FIGURE 1

Schematic views of the device.

$Ga_{0.47}$ $In_{0.53}$ As layer (about 1µm thick) is deposited on the undoped GaAs layer in order

to fabricate the photoconductive detector. This strained heteroepitaxy is grown by L.P.M.O.C.V.D.. Because of the higher refractive index of the GaInAs top layer, the propagation modes inside the GaAs optical waveguide become leaky and the light can be detected by the GaInAs photodetector. Two ohmic contacts are deposited on the GaInAs layer to fabricate the electrodes of the photoconductor. The electrode spacing is 50µm. The GaInAs layer has been etched (ion milling) down to the undoped GaAs layer except for the area corresponding to the photoconductive detector. A photograph of the device is given in figure 2.

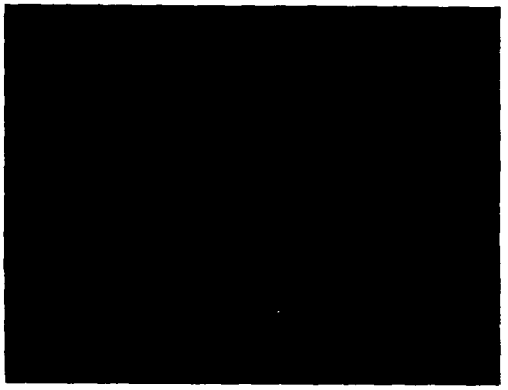


FIGURE 2

Photograph of the device.

The capabilities of the device to detect a part of the light propagating inside the waveguide have been tested. First, we present, in figure 3, the steady state photocurrent when the light (supplied by a 1.06µm YAG laser) is focused, via a microscope objective, on a cleaved edge of the waveguide.

By the comparison between these values and those obtained when the light is impinging on the top of the photoconductor, the ratio between detected light power and propagating light power has been found close to 10%. The influence of the coupling between the light beam and the cleaved edge of the waveguide has

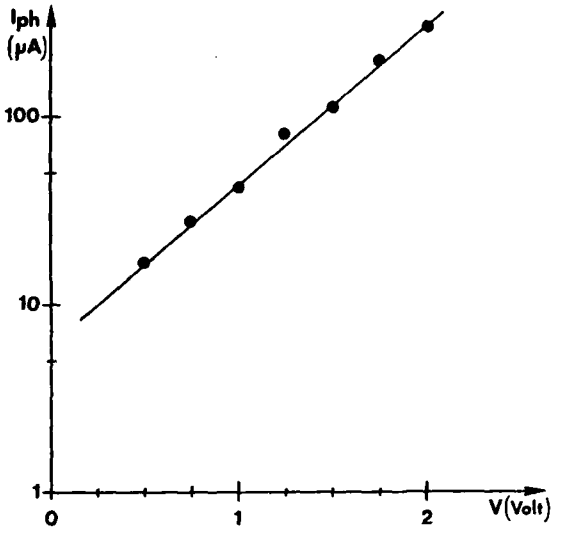


FIGURE 3

Measured photocurrent when light is propagating inside the waveguide (P1=220µW).

also been studied. As an example, we report the measured photocurrent when the light beam is shifted along an axis perpendicular to the epitaxial plane (figure 4); the highest photocurrent is observed for the maximum coupling.

A parameter which characterizes the photoconductor is the gain [4]; it is the number of electrons collected in the external circuit for one incident photon. In figure 5, we show the gain of the photoconductor versus bias voltage when the light is impinging on the top.

It can be observed that the gain increases when the bias voltage increases, corresponding to a decrease of the transit time τ_t of the electrons, according to the expression of the gain $G = \tau_v / \tau_v$ [5] [6] . τ_v is the electron-hole pair lifetime whose value could be connected to trapping effects at the $Ga_{0.47}$ $In_{0.53}$ As/ GaAs interface. The low value of the gain, as it has already been observed [7], is explained by a short lifetime τ_v, as it can be observed on the picosecond response (figure 6). This result indicates that this

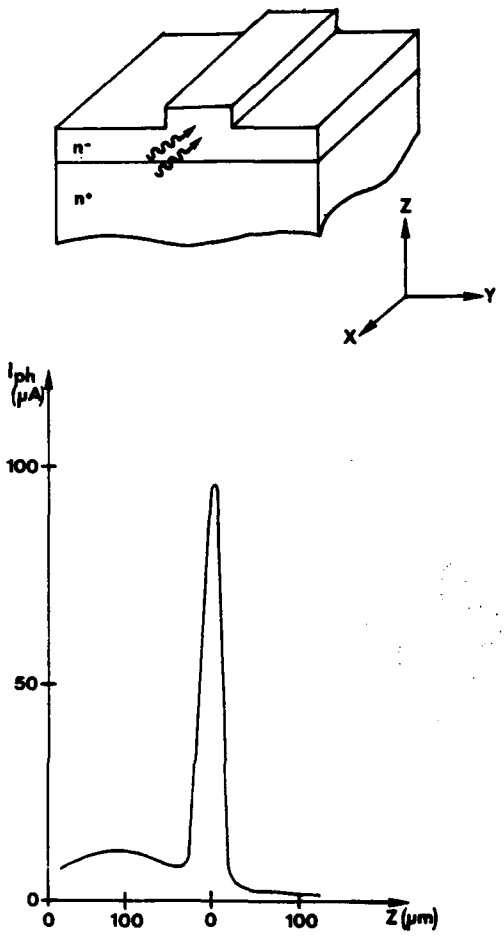


FIGURE 4

Measured photocurrent when the light beam is shifted along an axis perpendicular to the epitaxial plane.

detector could be useful for gigahertz receiver applications.

The noise properties of the photoconductor have been investigated in the 10MHz-1.5GHz frequency range using a HP 8970A noise figure meter. Our experimental results (figure 7) show a high 1/f noise for frequencies lower than 100MHz. As it is commonly observed in III-V materials [8] it is reduced for higher frequencies, to be close to the thermal (Nyquist) noise.

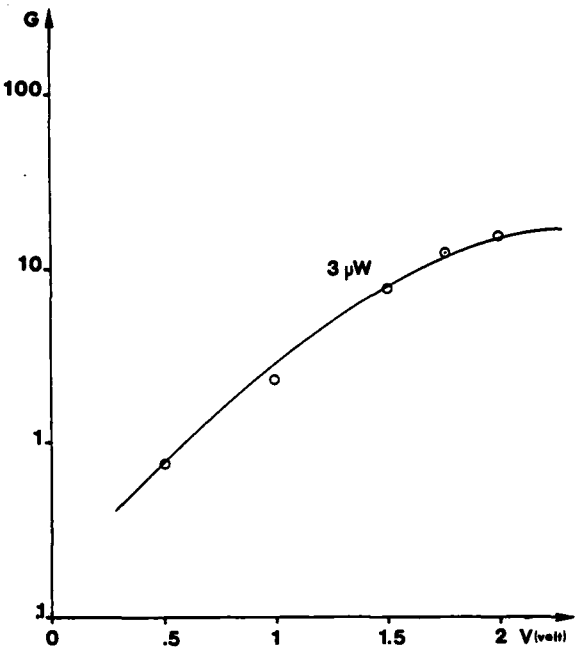


FIGURE 5

Gain of the photoconductor versus bias voltage when the light is impinging on the top.

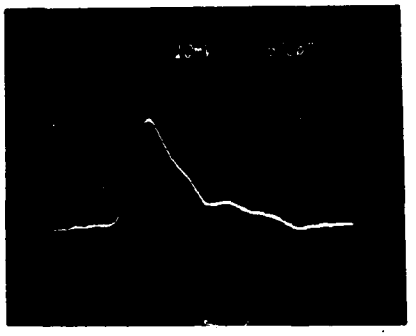


FIGURE 6

Picosecond response of the photoconductor.

In conclusion, our results show that it is possible to fabricate photoconductors monolithically integrated with rib waveguides suitable for 1.3-1.55μm wavelength operations using GaInAs/GaAs strained heteroepitaxies.

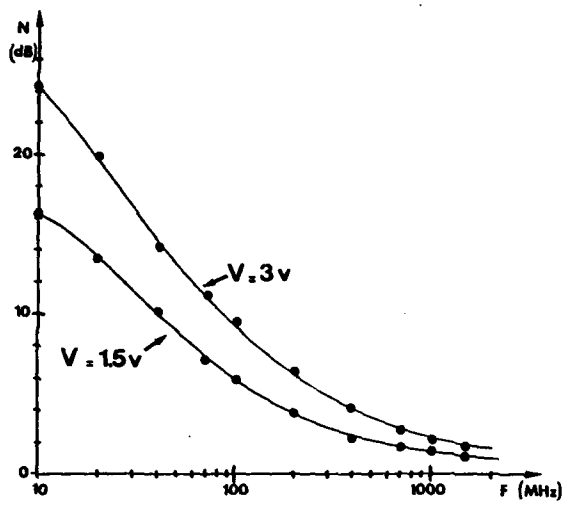


FIGURE 7

Noise measurements in the 10MHz-1.5GHz frequency range. (0 dB for 50Ω at 293 K.)

REFERENCES

[1] Bowers, J.E. Burrus, C.A., Elect. Lett. (1986) 22(17) pp.905-906.

[2] Chandrasekhar, S. Campbell, J.C. Dentai, A.G. QUA, J.C., Elect.Lett. (1987) 23(10) pp.501-502.

[3] Cinguino, P. Genova, F. Rigo, Cacciatore, C. and Stano, A., Appl.Phys. Lett., (1987) 50(21) pp.1515-1517.

[4] Chen, C.Y. Kasper, B.L. Cox, H.M., Appl. Phys.Lett., (1984) 44(12) pp.1142-1144.

[5] Beneking, H. I.E.E.E., (1982) E.D. 29 n°9 pp.1431-1441.

[6] Sze, S.M., Physics of Semiconductor Devices (Wiley & Sons, New-York, 1981).

[7] Razeghi, M. Ramdani, J. Verriele, H. Decoster, D. Constant, M. and Vanbremeersch, J., Appl.Phys.Lett, (1986), 49(4) pp.215-217.

[8] Vilcot, J.P. Constant, M. Decoster, D. and Fauquembergue, R. Physica B 129, 488 (1985).

MBE GROWTH AND PROCESSING OF InGaAs/InGaAlAs/InP MONOLITHICAL INTEGRATED RIDGE WAVEGUIDE PHOTODIODES

Piero CINGUINO, Fernando GENOVA, Cesare RIGO, Carmelo CACCIATORE and Alessandro STANO

CSELT - Centro Studi e Laboratori Telecomunicazioni S.p.A. -
Via G.Reiss Romoli, 274 - 10148 Torino (ITALY)

The molecular beam epitaxial growth of high quality InGaAlAs/InP and its application to low loss passive optical waveguides and waveguide-integrated photodiodes for operation at 1.55 μm is described. The chemical etching characteristics of the layers were optimized in order to fabricate low loss (2.2 dB/cm) ridge waveguides. Integration of the waveguides with an InGaAs pin photodiode is demonstrated using absorption of the guided light by leaky coupling from the InGaAlAs guiding layer into the higher index InGaAs absorbing region containing a p/n junction. The devices showed external quantum efficiencies as high as 20% for operation at 1.55 μm wavelength. This is the first demonstration of a monolithic integrated waveguide device in the InGaAlAs/InP material system.

1. INTRODUCTION

Monolithic integration of optical and optoeletronic functions on a single chip is a strong requirement for the development of lightwave communication systems at the wavelengths of 1.3 and 1.55 μm [1]. A basic component for such circuits is a photodiode directly integrated with an interconnecting low loss waveguide to provide either an end-line or a monitor detector. Several applications are, for example, laser-monitor detector integration, end-line detection in wavelength demultiplexing or in heterodyne optical circuits, in-line monitor detection for feedback operations. Here we report on MBE growth of high quality InGaAlAs /InP and its application to low loss optical waveguides and monolithical integrated waveguides photodiodes for operation at 1.55 μm.

2. InGaAlAs MBE GROWTH

High quality InGaAlAs/InP layers with bandgap 0.85μm (InAlAs), 1.042μm, 1.29μm and 1.63μm (InGaAs) were grown in a VG80H twin chamber MBE system with continuous substrate rotation, provided with an interlock for the fast reloading of the arsenic contained in a cracker cell. The determination of the proper growth conditions was obtained optimising the intensity of photoluminescence spectra and their full-width at half maximum (FWHM) as described elsewere [2]. The composition with bandgap 1.29 μm used in devices fabrication showed defect densities as low as 2000 cm^{-2}, with a 4K full width at half-maximum (FWHM) of 23 meV and double crystal (DC) x-ray profiles with FWHM of 17.5" of arc, with background doping levels as low as 3 10^{15} cm^{-3} [2].

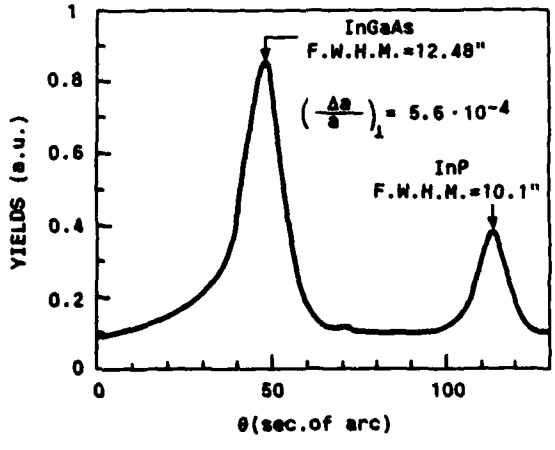


Fig. 1 - X ray DC profile for a ternary InGaAs layer

InGaAs/InP heterostructures showed a 4K photoluminescence FWHM of 6.1 meV on 2.5 μm thick n-doped samples,with a DC x-ray FWHM of

12.48" of arc as shown in Fig. 1, which is the narrowest linewidth ever reported for this material [2].

3. DEVICE TECHNOLOGY AND RESULTS

To test waveguiding in InGaAlAs/InP at 1.55 μm, ridge waveguides were chemically etched on quaternary layers with bandgap 1.2 μm. A detailed study of the chemical etching characteristics of (001) oriented InAlAs/InP and InGaAs/InP heterostructures was previously carried out [3], in order to define the best conditions for waveguide morphology and to calibrate the etch rates for devices fabrication. The etch rates and the etch-revealed planes were determined on stripes oriented along the [110] and [1$\bar{1}$0] directions and on circular mesa structures for the H_3PO_4 : H_2O_2 : H_2O, H_2SO_4 : H_2O_2 : H_2O and Br_2 - CH_3COOH etching systems. Similar etching characteristics were obtained for the InGaAlAs/InP quaternary layers. Losses as low as 2.2 dB/cm at 1.55 μm were observed on ridge structures grown both on n+ and semi-insulating InP substrates [4]. Low loss waveguiding (<5 dB/cm) at 1.3 and 1.55 μm was also observed on ridges made on the quaternary composition with bandgap 1.042 μm. To integrate the ridge waveguides with an InGaAs photodiode we used the structure schematically illustrated in Fig.2.

Absorption of the guided light is provided by leaky coupling from the quaternary guiding layer into the higher index InGaAs absorbing layer. The use of the ridge structure has the advantage of not requiring regrowth onto non-planar substrated or localized growth of the absorbing InGaAs material. The layers sequence consists of a 2.5 μm thick undoped InGaAlAs layer with bandgap 1.29 μm, followed by an undoped InGaAlAs layer with a thickness of .5 μm and then by a 2.5 μm thick p+ InGaAs layer.

The device structure was defined by first etching the photodiode mesa through a Si_3N_4 mask down to the InGaAs/InGaAlAs interface with a H_3PO_4 : H_2O_2 : H_2O 1:8:40 etching solution. Then the ridge waveguides were etched through a photoresist mask with a ridge step of 1 μm. Finally, the top contact geometry was defined, and the structure was passivated with photo-CVD Si_3N_4 [5]. The top contact was made in an external lateral pad, in order to bond

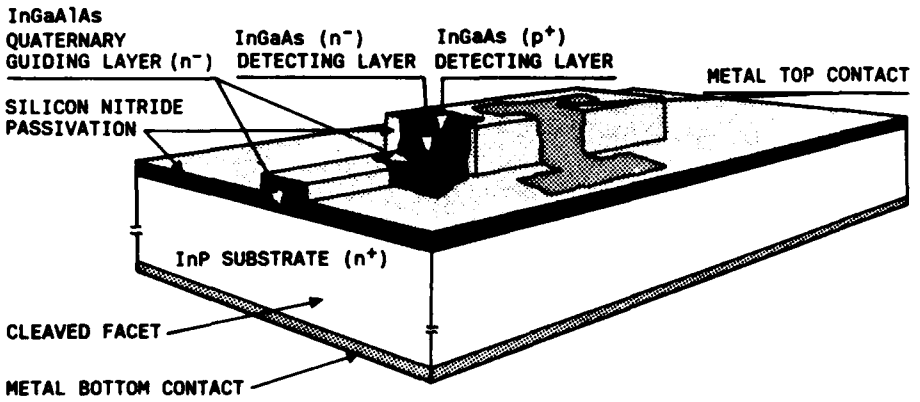


Fig. 2 - Schematic layout of the waveguide integrated with the PIN detector

the devices outside the photodiode mesa. A scanning electron micrograph of a mounted device is shown in Fig. 3, showing the lateral contact pad geometry and the cleaved facets of the ridge waveguide.



Fig. 3 - Micrography of a mounted device

Ti/Au and Au-Sn were used as top and bottom contacts, respectively. The device lengths were 200, 300 and 400 μm, while the ridge widths were 10, 20, 30 and 50 μm. The measured external quantum efficiencies of the devices are as high as 20% for butt-coupling to a single mode fiber. After correction for reflection, coupling and propagation losses, the devices internal quantum efficiencies are estimated to be 80-100%, depending on the device length. Bandwidths of 600 MHz have also been measured, which can be greatly increased by reducing the photodiode dimensions to obtain low capacitance diodes for high speed operation [6].

4. CONCLUSIONS

In conclusion, we have demonstrated the first integrated waveguide device in the InGaAlAs/InP material system by integrating a low loss InGaAlAs ridge waveguide with an InGaAs photodiode for operation at 1.55μm. This demonstrate the potential of InGaAlAs for the development of high quality waveguide devices for integrated optics applications.

ACKNOWLEDGEMENTS

This work has been supported by the European Economic Community under the ESPRIT 263A Project.

REFERENCES

[1] Wada O., Sakurai T., and Nakagami T., IEEE J. Quantum Electron.QE-22 (1986) 805

[2] Genova F., Morello G., and Rigo C., in: Proc. 7th MBE Workshop (1986), to be published in J.Vac.Sci.Technol.B.

[3] Stano A., J.Electrochem.Soc., 134 (1987) 448

[4] Cinguino P., Genova F., Morasca S., Rigo C. and Stano A., Electron.Lett. 23 (1987) 235 and in: Wilkinson C.D.W. and Lamb J., (eds.), Proc. ECIO'87 (SETG, Glasgow, 1987), pp. 78-81

[5] Meliga M., Stano A., and Tamagno S., in: "Thin Films Technologies", Proc. SPIE, Vol. 652 (SPIE, Bellingham, 1986) pp.243-247

[6] Bowers J.E. and Burrus C.A., Electron. Lett. 22 (1986) 905

Session B4.2

Bipolar Modelling II

Chairman: H.C. De Graaff

Thursday, September 17, 1987

A New Simple Analytical Evaluation of Minority Carrier Current in Arbitrarily Doped Region with Non-thermal Generation of Carriers

C. R. Selvakumar and D. J. Roulston

Deptartment of Electrical Engineering, University of Waterloo, Waterloo, Canada N2L 3G1

A simple general analytical solution to the minority carrier transport equations in an arbitrarily doped semiconductor (Si, GaAs, and InGaAsP) region is obtained by including an arbitrary non-thermal source term in the continuity equation. Internal quantum efficiencies resulting from AM1 illumination in gaussian doped silicon emitters are calculated by the new analytical expression and compared with "exact" computer calculations.

1. Introduction.

Evaluation of minority carrier current in a semiconductor region in the presence of external generation is a classical problem. When the semiconductor region wherein generation takes place has doping gradients and has been moderately or heavily doped (as for example in emitters of solar cells), an analytical evaluation of minority carrier current in such a region becomes extremely difficult and invariably one resorts to numerical methods of solution because the lifetime, mobility and bandgap become position dependent.

Taking into account experimentally determined doping dependencies of lifetime (SRH and Auger), bandgap narrowing and mobility we have obtained for the first time a simple and yet a general *analytical* solution to minority carrier transport equations in an arbitrarily doped semiconductor (Si, GaAs and InGaAsP) region by *including* an *arbitrary*, non-thermal generation of carriers such as due to light, e-beam, x-rays etc. This solution is useful for numerous applications in solar cells, photo detectors etc.

2. Theory

A *new relation between transport parameters* is recognized [1] and is used in deriving the analytical solution; the new relation being

$$\frac{D_m m_o}{L_m} = C_s = \text{constant} \qquad (1)$$

where D_m is minority carrier diffusion coefficient ($m = n$ or p to denote electrons or holes), m_o is thermal equilibrium minority carrier density and L_m is minority carrier diffusion length in the semiconductor. Recently Del Alamo and Swanson [2] have experimentally measured the product $p_o D_p$ and diffusion length L_p in N-type silicon as functions of doping. We have plotted the ratio $D_p p_o / L_p$ as a function of doping in Fig. 1 and we find this ratio to be essentially a constant [1].

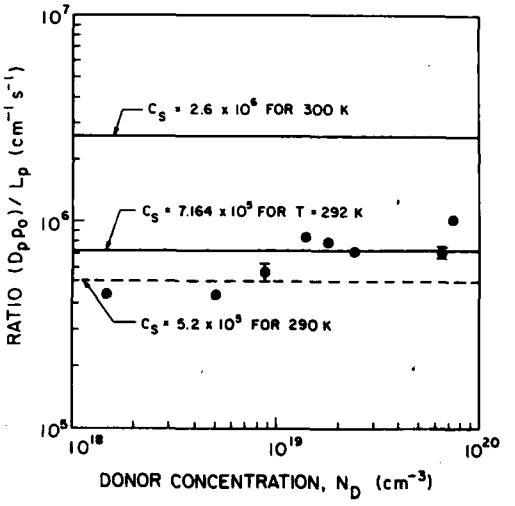


FIGURE 1

The newly defined quantity $D_p p_0/L_p$ plotted as a function of donor density N_D in silicon . The product $D_p p_0$ and L_p are experimentally determined by Del Alamo and Swanson [6]. The lines are from references 4 and 5.

The experimental data of Del Alamo and Swanson [2] used in calculating this ratio in (1) were obtained at 292°K. The value of C_s at 300°K plotted in Fig. 1.

was determined by fitting independent lifetime data of other workers as shown in [4-5] and the value of C_e at 292°K plotted in Fig. 1 was determined merely by interpolation. As seen in Fig. 1, this value of C_e (292°K) is found to be a remarkably good fit of the relation (1).

This relation had been found to be true for GaAs and InGaAsP as well [3-4]. Using (1) we obtain minority carrier transport equations in terms of the normalized excess minority carrier density $u(=m/m_o$; m is excess minority carrier density) as follows

$$|J_m| = qC_eL_m\frac{du}{dx} \tag{2}$$

$$\frac{d^2u}{dx^2} + \frac{1}{L_m}\frac{dL_m}{dx}\frac{du}{dx} - \frac{u}{L_m^2} = -\frac{G(x)}{C_eL_m} \tag{3}$$

The general solution of the continuity equation (3) using the standard boundary conditions given in (12) and (13) is

$$u(x)=[A(x) + A_1]\sinh Ki(x)+ [B(x) + B_1]\cosh Ki(x) \tag{4}$$

and therefore

$$|J_m(x)|=qC_e[\{A(x)+A_1\}\cosh Ki(x)+ \{B(x)+B_1\}\sinh Ki(x)] \tag{5}$$

where $A(x)$ and $B(x)$ are simple integrals involving generation rate $G(x)$ and one hyperbolic function of $Ki(x)$ and A_1 and B_1 are constants as follows:

$$A(x) = -\frac{1}{C_e}\int_0^x G(\lambda)\cosh Ki(\lambda)d\lambda \tag{6}$$

$$B(x) = \frac{1}{C_e}\int_0^x G(\lambda)\sinh Ki(\lambda)d\lambda \tag{7}$$

$$A_1 = -\frac{A(w)\sinh Ki(w)+B(w)\cosh Ki(w)}{\sinh Ki(w)+S_F\cosh Ki(w)} \tag{8}$$

$$B_1=S_FA_1 \tag{9}$$

and the two *dimensionless* parameters Ki number and Surface Factor S_F are given by

$$Ki(x) = \int_0^x \frac{d\lambda}{L_m(\lambda)} \tag{10}$$

$$S_F = \frac{D_m(0)}{S_mL_m(0)} \tag{11}$$

The boundary conditions are

$$u(w) = 0 \tag{12}$$

$$J_m(0) = qS_m m_o(0)u(0) \tag{13}$$

where S_m is the recombination velocity of minority carriers at the contact ($x = 0$) and the junction is assumed to be at $x = w$.

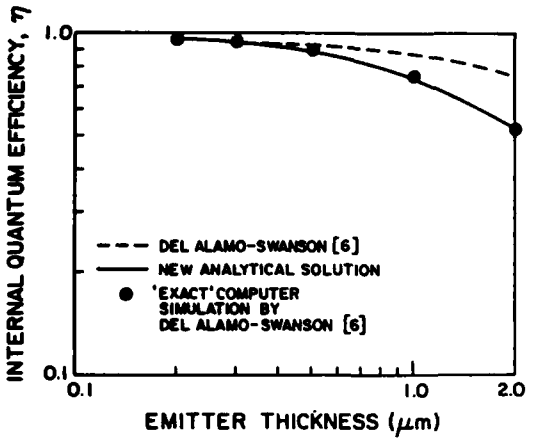


FIGURE 2

Comparison of the new analytical evaluation of internal quantum efficiency with the "exact" computer solutions of Ref. 6. The dashed curve is due to the analytical method of Del Alamo and Swanson in Ref. 6. Silicon gaussian emitters with surface doping density $N_D = 10^{20}\text{cm}^{-3}$ and background p-type doping density of $N_b = 10^{16}\text{cm}^{-3}$ are assumed. Surface recombination velocity for holes $S_p = 10^4\text{cm/s}$.

3. Results and Discussion

Using the analytical expression in (5) we have calculated the internal quantum efficiency of Gaussian doped n-type silicon emitters for AM1 illumination using the rates given in [7]. Our results are compared with the "exact" computer calculations of Del Alamo and Swanson [6] in Fig. 2 and 3 for different emitter thicknesses.

We have assumed that temperature T = 300°K and $C_e = 3.16 \times 10^6\ \text{cm}^{-1}\text{s}^{-1}$ [8] for all the comparisons in Figures 2 and 3.

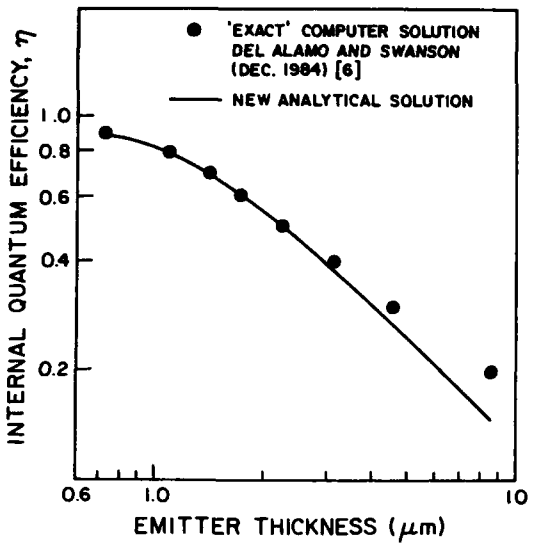


FIGURE 3

Comparison of the present analytical evaluation of internal quantum efficiency with the exact computer simulation results of Ref. 6. for thicker emitters. Surface recombination velocity for holes $S_p = 10^3$ cm/s and other parameters are the same as for Fig. 2.

As we can see from these figures, the predictions of the simple expression in (5) are remarkably close to the exact computer solution even for *thick* emitters. Appreciable errors occur only for Si emitters thicker than 5 μm; even then acceptable errors of around 10 - 20% are encountered which are well within the accuracy with which we can determine the transport parameters.

References

[1] C.R. Selvakumar, private communication to R.M. Swanson dated Dec. 13, 85.

[2] J.A. Del Alamo and R.M. Swanson, Record of the 18th IEEE Photovoltic Spec. Conf. (1985).

[3] C.R. Selvakumar, Solid-St-Electronics (in press).

[4] C.R. Selvakumar, Ph.D. Thesis, Indian Institute of Technology, Madras, India (1984).

[5] C.R. Selvakumar. J. Appl. Phys. 58, 3476-3478 (1984).

[6] J.A. Del Alamo and R. M. Swanson, IEEE Trans. Electron Devices, ED-31, 1878-1888 (1984).

[7] J. Furlan and S. Amon, Solid-St-Electron, 28, 1241-1243 (1985).

[8] C.R. Selvakumar and D. J. Roulston, Solid-St. Electronics (in press).

MEASUREMENT AND CALCULATION OF BASE-RESISTANCE COMPONENTS OF MODERN HIGH-SPEED BIPOLAR TRANSISTORS

J. Fertsch, H. Voit, H. Klose, and W.R. Böhm

Siemens AG, Central Research and Development
Otto-Hahn-Ring 6, D 8000 München 83, FRG

A numerical method based on the Gummel-Poon model is presented, which permits to calculate the components of the base resistance from the layout and easily accessible DC-parameters. The calculated results are compared with results derived from measured S-parameters.

1. INTRODUCTION

One of the most important factors limiting the switching speed of high-speed transistors is the base resistance, since the base has to be charged or discharged via this resistance. Thus, its exact knowledge is indispensable for optimizing the circuit design. Conventional analytical methods to estimate the base resistance are no longer sufficient because of second order effects such as current crowding, base width modulation, Early- and Kirk-effect. Our method is based on the Gummel-Poon model and hence includes these effects, automatically.

2. CALCULATION

Fig.1 shows a schematical cross-section of the emitter-base complex of an advanced self-aligned bipolar transistor with polysilicon contacts (PSA).
The four components of the base resistance are:

1) contact resistance of the base
2) resistance of the polysilicon on SiO_2
3) resistance of the polysilicon / monosilicon-diffused double layer
4) the resistance of the active base region

No linkage problem between region (3) and (4) was observed in our transistors. Therefore this division is a complete one.

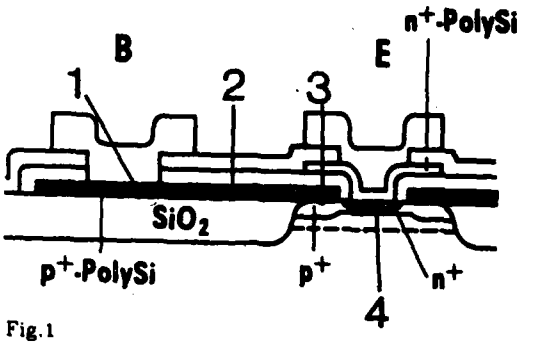


Fig.1

Cross section through the emitter base complex showing the four components of the base resistance.

The base-resistance calculation which is three-dimensional in reality, was solved in two dimensions in the top view projection. This approximation is certainly permissible if the thickness of the layer is small enough. Even in the double layer region (3) a two-dimensional treatment is justified if the ratio of the length l to the thickness t of the double layer satisfies $\frac{l}{t} > 5$ [1].

After transformation to two dimensions, the following continuity equations are obtained for the four components [1].

$$1)\ \nabla\cdot\left(\frac{1}{R_{SO}}\cdot\nabla\psi\right) = -\frac{U_{BE}-\psi}{\rho_c}$$

$$2)\ \nabla\cdot\left(\frac{1}{R_{SO}}\cdot\nabla\psi\right) = 0$$

$$3)\ \nabla\cdot\left(\frac{1}{R_{SM}}\cdot\nabla\psi\right) = 0$$

$$4)\ \nabla\cdot\left(\frac{QB(\psi)}{R_{SP}}\cdot\nabla\psi\right) = \frac{I_S}{A_E\cdot\beta_0}\cdot\left[\exp\left(\frac{e\psi}{kT}\right)-1\right]$$

The meaning of the used symbols is:

ψ potential within the semiconducting layer
R_{SO} sheet resistance of poly-Si on oxide
R_{SM} sheet resistance of poly-Si on mono-Si
R_{SP} sheet resistance of the active base
ρ_c specific contact resistivity
QB normalized Gummel-charge
The other symbols have the usual meaning.

Equation (4) is based on the Gummel-Poon model. These four equations were solved simultaneously with a numerical treatment, taking into account the following boundary condition:

$$\left.\frac{\partial\psi}{\partial n}\right|_{\Gamma} = 0$$

A typical solution for the potential is shown in fig.2. It was calculated over only half the area for reasons of symmetry.

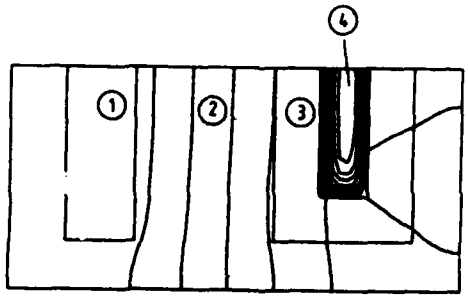


Fig.2

Equipotential lines within the base region at $U_{BE} = 860\ mV$.

It can be seen that at $U_{BE} = 860\ mV$ the largest voltage drop occurs over the active base region. At all sides of that region there is nearly the same current density. If $\psi(x,y)$ is known, then by means of the energy conservation law

$$R_b\cdot I_B^2 = \frac{1}{R_{SO}}\int_{(1)}(\nabla\psi)^2\,dx\,dy + \int_{(1)}\psi\cdot\frac{U_{BE}-\psi}{\rho_c}\,dx\,dy + \frac{1}{R_{SO}}\int_{(2)}(\nabla\psi)^2\,dx\,dy + \frac{1}{R_{SM}}\int_{(3)}(\nabla\psi)^2\,dx\,dy + \frac{1}{R_{SP}}\int_{(4)}QB(\psi)\cdot(\nabla\psi)^2\,dx\,dy$$

the partial resistances can be calculated for each individual region.

3. MEASUREMENT

The base resistance was determined experimentally by using a network analyzer to measure the S-parameters at different collector currents in the frequency range from 50 MHz to 3 GHz. The measurements were performed directly on the wafer. This greatly simplifies the measurements since no time-consuming mounting with inherent impedance mismatch is necessary, and hence improves considerably the calibration accuracy. The input impedance Z_{in} was calculated from the S-parameters [3]

$$Z_{in} = \frac{(1+S_{11})\cdot(1+S_{22}) - S_{12}\cdot S_{21}}{(1-S_{11})\cdot(1+S_{22}) + S_{12}\cdot S_{21}}\cdot Z_0$$

In Fig.3 the imaginary part of Z_{in} was plotted against its real part. Up to a frequency of 1GHz, the measured data can be described very well by a semicircle in the complex plane showing that the simple 1-pole approximation [2] is valid. Above 1 GHz deviations from the semicircle are observed. The small signal base resistance r_b was determined by means of curve fitting and then converted to the large signal base resistance R_b by suitable averaging

$$R_b = \frac{1}{I_c}\int_0^{I_c} r_b(I_c')\,dI_c'$$

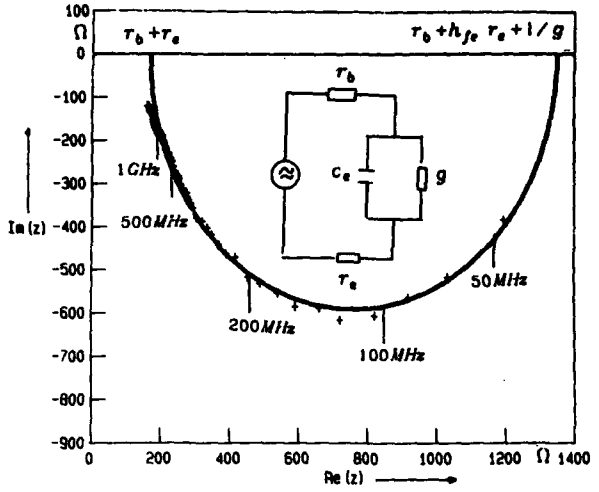


Fig.3

Imaginary part of input impedance $Im(Z_{in})$ versus $Re(Z_{in})$.
The measured values are marked by crosses.
The fitting curve (solid line) is a semicircle in the lower half Z-plane.
The smaller zero position is the sum of the small signal base resistance r_b and emitter resistance r_e.

4. RESULTS

Fig.4a and 4b show the calculated total base resistance R_b and its external component $R_{b\,ex}$ as a function of the collector current density j_c. A comparison was made for transistors with emitter mask areas $2\times8\mu m^2$ and $2\times40\mu m^2$, with one or two base contacts respectively.

It is of interest to note that the relative reduction of the resistance was by a factor of about 2 in the case of long emitter transistor and by a factor of only 1.3 in the case of the short one. The experimental results are also plotted in Fig.4. The good agreement obtained between the measured and calculated results can be seen as a confirmation of the method of calculation used.

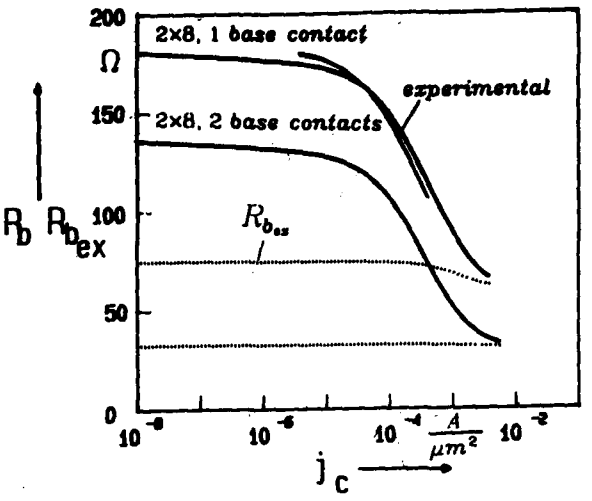


Fig.4a

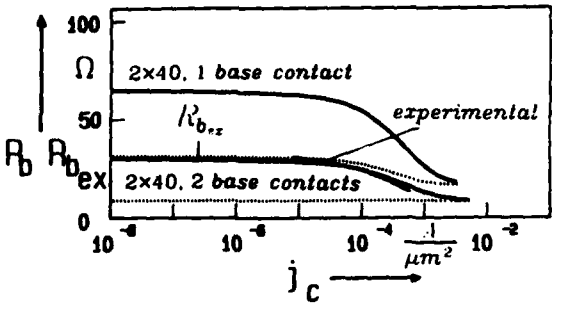


Fig.4b

Fig.4a and 4b

Calculated base resistance (solid line) and its external component (dotted line) vs. collector current density. Experimental results are also shown.
a) $A_E = 2\times8\ \mu m^2$
b) $A_E = 2\times40\ \mu m^2$

5. CONCLUSION

In this paper, a numerical method was shown to determine the components of the base resistance, especially the external and internal base resistance, as a function of the collector current density. The task consisted essentially in solving the stationary continuity equation within the base region under the two-dimensional approximation. Further, the base resistance was determined experimentally with the aid of the input impedance circle diagram method from measured S-parameters. Experimentally and numerically determined base resistances show good agreement. The methods used are accordingly correct. The advantage of the calculation is that the components of the base resistance can be determined at any arbitrary collector current, whereas the base resistance as a whole can be measured only in a small current range.

REFERENCES

[1] Josef Fertsch: Physikalische Untersuchungen an selbstjustierten Bipolartransistoren-Messung und Berechnung der Anteile des Basisbahnwiderstands, Diplomarbeit Univ. Regensburg 1986.

[2] Ian Getreu: Modeling the Bipolar Transistor, Elsevier Scientific Publishing Company 1978.

[3] J. Lehmann: Dioden und Transistoren, Vogel-Verlag, p.94, 3. Auflage, 1972, Würzburg.

A BIPOLAR DC MODEL FOR TRANSISTORS FABRICATED IN A CMOS PROCESS

Denis J. F. Doyle, William A. Lane

National Microelectronics Research Centre,
University College, Cork,
Ireland.

Abstract - an isolated vertical npn transistor fabricated in an n-well CMOS process is described. Characterisitic features of the transistor are examined, particularly in relation to the absence of a buried layer and low well doping. An equivalent circuit is presented to model the structure as a 4-terminal device, which can be implemented in SPICE without modification to the SPICE BJT model. A parameter extraction sequence for the model is detailed and the results of a parameter optimisation for a real device are presented.

1. INTRODUCTION

Within many CMOS processes the formation of a junction isolated bipolar transistor is possible without the addition of extra masking steps [1]. Because of the restrictions imposed by CMOS processing, this bipolar device exhibits features which differ from those fabricated using conventional bipolar processes. Some of these features place limitations (1) on the device usuage and (2) on the ability of the SPICE BJT model to predict any unusual circuit behaviour. It is necessary that designers have the capability to model all circuit conditions, for example, the high substrate current that results if excess base drive is applied or during circuit power-up.

In this paper a junction isolated npn transistor fabricated using a 5 µm, n-well CMOS process is described. An equivalent circuit is presented, for the structure, which can be fully implemented in SPICE without changing the BJT model. This equivalent circuit models the effects of (1) the vertical parasitic substrate pnp transistor, (2) the high-value collector resistance and (3) the JFET-like action of the collector region. A parameter optimisation sequence is described for obtaining all equivalent circuit parameters from conventional bipolar measurements and the results of this optimisation for a real device are presented.

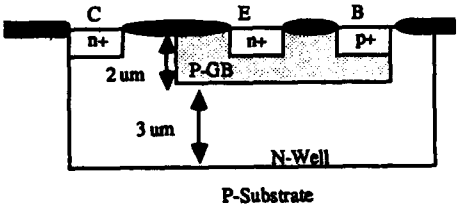


Fig. 1: Isolated NPN Structure

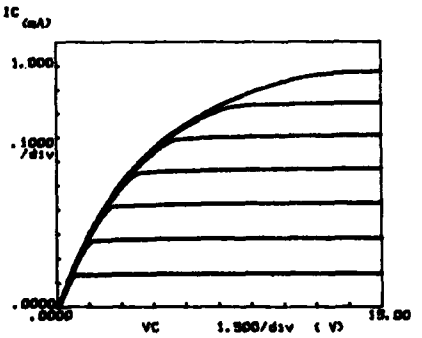


Fig. 2; NPN Forward Characteristic

2. NPN DEVICE STRUCTURE

The structure of the bipolar is shown in fig. 1. The CMOS n-well region forms the collector, the P+ guard band implant forms the active base region and the N+ source/drain implant forms the emitter. The total well depth is 5 µm, base region depth is 2 µm and the active base width is approximately 1 µm. The typical gain of the transistor is 130, Early Voltage is 100V and the breakdown voltages are as follows: BV_{ebo} = 10V, BV_{cbo} = 55V and BV_{ceo} = 55V. A typical forward characteristic for this structure, with an

emitter length of 26 μm and an emitter width of 13μm is shown in fig. 2.

3. BIPOLAR STRUCTURE MODEL

The most notable feature of the device structure in fig. 1 is the absence of an collector N+ buried layer. This, together with the low doping concentration in the n-well and the well depth, has a number of effects on the device behaviour. Firstly, the parasitic vertical pnp transistor, formed by the base, collector and substrate of the npn device, has a high forward current gain (Bf_{pnp}) of typically 100. The connection of this pnp is shown in fig. 3. This parasitic transistor does not affect the operation of the npn device when it is in the linear active region, as is the case in most analog circuits, since the base/collector junction of the npn transistor is reverse biased, as is the collector/substrate junction. This effectively turns off the pnp device. However, when the npn device is saturated, both the base/emitter and base/collector junctions of the npn are forward biased, placing the pnp device in its forward linear region. This leads to current flow into the substrate, a situation that should be avoided in CMOS.

The approach used to model the current flow to the substrate is to ensure that in the npn collector current equation, any terms which represent hole injection into the n-type collector region (where they are now minority carriers in the pnp base) are transferred to the substrate with an efficiency ($1-1/Bf_{pnp}$). The npn collector current equation is as follows:

$$\underset{\text{(i)}}{} \qquad \underset{\text{(ii)}}{}$$

$$Ic = \frac{Is}{Qb}\left[\exp\left(\frac{Vbe}{n_f V_T}\right) - \exp\left(\frac{Vbc}{n_r V_T}\right)\right]$$

$$- \frac{Is}{Br}\left[\exp\left(\frac{Vbc}{n_r V_T}\right) - 1\right] \qquad \text{(iii)}$$

$$- Isc\left[\exp\left(\frac{Vbc}{n_c V_T}\right) - 1\right] \qquad \text{(iv)}$$

(with the usual SPICE parameter notation)

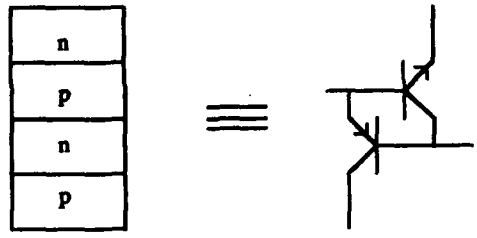


Fig. 3: NPNP Structure

Terms (i) and (ii) represent the collector electron (or minority carrier) current across the base region under forward active and reverse active conditions respectively. Terms (iii) and (iv) are the base currents associated with the base/collector junction and are in fact hole current terms. The collector current equation therefore consists of the electron current minus the terms which represent hole currents injected in to the collector region [2]. However, as pointed out above, these hole currents constitute minority carriers in the n-type collector region and are therefore collected by the reverse biased collector/substrate junction, with efficiency ($1-1/Bf_{pnp}$) and only a fraction $1/Bf_{pnp}$ of terms (iii) and (iv) should remain in the collector current equation.

Fortunately, the correct equations for the current flow in the 4-terminal structure can be implemented as an equivalent circuit, without any modification to the BJT model. This is achieved by replacing terms (iii) and (iv) above with two pnp transistors, the first having parameters Is_{pnp} = Is/Br and nf_{pnp} = nr, the second having Is_{pnp} = Isc and nf_{pnp} = nc. Both transistors have the same beta = Bf_{pnp}. These terms are then eliminated from the npn equations by specifying Br = 1.0E06 and Isc = 0. This results in the transfer of hole current to the substrate and the correct terminal collector current being supplied from the electron current terms minus the pnps' base currents [3].

The second important effect of the absence of the buried layer is on the current path between collector contact and the point of current entry to the intrinsic npn device. The resistance of

this path is high and nonlinear. This leads to a rapid gain roll-off with collector current, low current saturation of the transistor, a high Vce_{sat} and a limited frequency response [4]. This current path can be seen to be made up of two distinct regions as in fig. 4. The first of these, Rc_1, is the region connecting the collector terminal to the base implant edge. This is modelled by a constant resistance. The value of this resistance is set by the well doping and the design rule spacing. The second path in the series, Rc_2, is the region of well beneath the base implant. This region is 3 μm deep for the structure in fig. 1 and its resistance is highly dependent on the base/collector voltage due to depletion spreading. In fact, the base region effectively acts as the gate of an n-channel JFET [5]. The value of the pinch-off voltage for this channel region depends on the channel thickness and doping. A typical value for the structure in fig. 1 is 15V. However, the effective pinch-off voltage may be much less than this depending on the substrate voltage which constitutes a back-gate bias for the channel. The effect of this nonlinear channel is clearly demonstrated in the forward characteristic of fig. 2, where the saturated/active region boundary rapidly rolls over into the active region with increasing base current, thereby limiting the collector current in the device.

The overall equivalent circuit is shown in fig. 5. Two important points need to be clarified for use of the JFET in the equivalent circuit. Firstly, only the JFET channel conduction equations are being used, as any gate current will be already modelled as npn base current. Secondly, the SPICE JFET model does not allow for back-gate bias and hence a SPICE MOSFET Level 1 model is used, which has the same conduction equations as a JFET.

4. PARAMETER EXTRACTION

Parameters are extracted for the complete equivalent circuit using standard bipolar

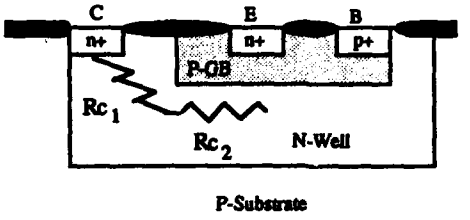


Fig. 4: Collector Resistance Components

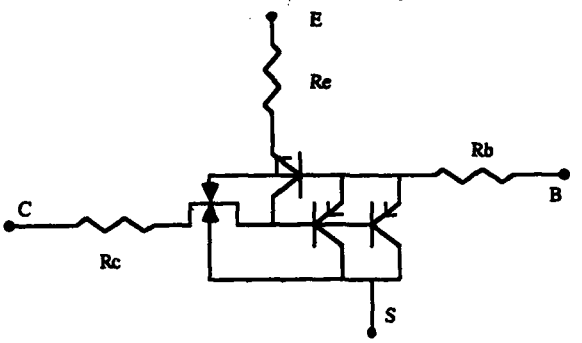


Fig. 5:Complete Equivalent Circuit

measurements. The parameters are optimised using a parameter optimisation program, which allows parameters to be fitted selectively to data regions where they are most relevant and in sequence [6]. The optimisation program also allows use of different error functions. In particular many of the optimisations involve fitting to the conductance, g_{ce}, (i.e. slope of the data) and transconductance, g_m, as well as purely to DC values. This is due to the importance of conductances and transconductances in small signal analysis of analog circuits. The optimisation sequence is summarised in Table 1.

The optimised model fits for a device as in fig. 1, are shown in fig. 6 and fig. 7 for the forward characteristic and Gummel plots, respectively. The current limiting effect is well modelled in fig. 6. This could not be modelled by considering the npn transistor solely. The overall average absolute DC and conductance errors are 5% and 22% respectively. The Gummel plots show not only the base and collector currents, but, also the emitter and substrate currents. The model, therefore, fits

Table 1: Parameter Optimisation Sequence

Stage	Data Region	Parameters	Error Function
1	linear active Gummel Plot	nf, ne	g_{ce}
2	linear active Gummel Plot	Is, Ise	DC
3	forward characteristic	Rc, Kp Bf,Gamma	DC, g_{ce}, g_m
4	linear active Gummel Plot	Ise, ne	DC, g_{ce}
5	complete Gummel Plot	Rb, nr, Br, Bf_{pnp}	DC, g_{ce}
6	forward characteristic	Vaf,Rc,Kp, Vto,lambda, Phi, Gamma	DC, g_{ce}

all currents in the device, whereas considering an npn transistor alone does not model substrate current. In addition, it is clear from the Gummel plots that no substrate current flows when the npn device is in its linear active region. Currents flow in the substrate commences only at the onset of saturation.

Some operating point analysis has to be performed for the paramater optimisation program to solve for equivalent circuit internal voltages, as only terminal voltages are accessed during device measurement. This is achieved by using a damped Newton convergence scheme to solve for node voltages.

5. CONCLUSIONS

In this paper a NPN bipolar transistor in a 5 µm, n-well, CMOS process was presented. The effects produced in such a device by the absence of a buried layer and the current limiting effects of a resulting JFET-like structure were discussed. An equivalent circuit was presented which treats the npn as a 4-terminal structure by modelling substrate current flow. Some parameter optimisation results from standard bipolar measurements were presented which showed good DC and conductance agreement between the equivalent circuit and measured results.

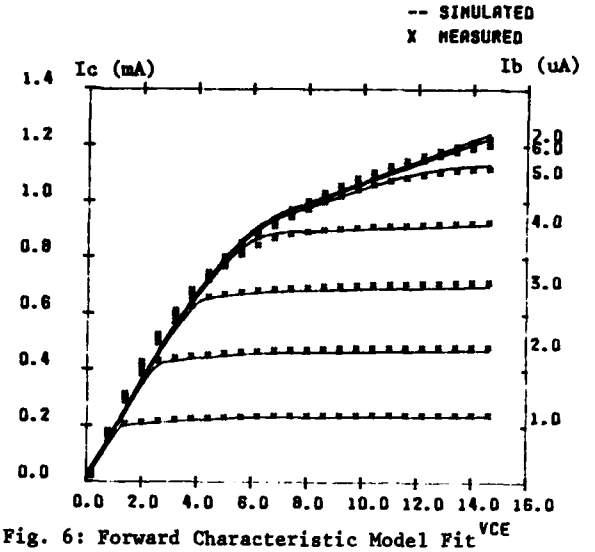


Fig. 6: Forward Characteristic Model Fit

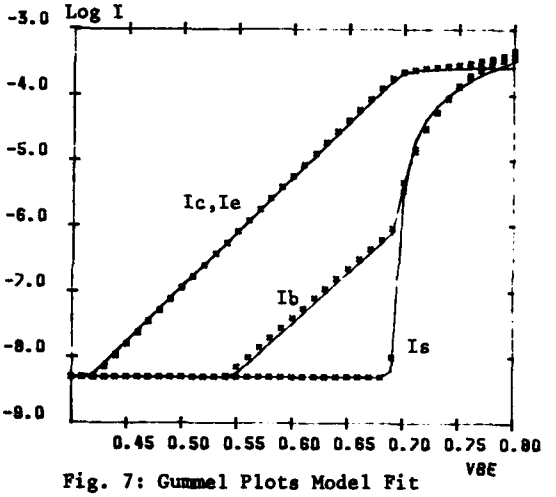


Fig. 7: Gummel Plots Model Fit

REFERENCES

[1] H. Monose, et al, IEEE Trans. on Electron Devices, Vol. ED-32, No. 2, February 1985, pp. 217-223.

[2] Ian Getreu, "Modeling the Bipolar Transistor", Tektronix, 1976.

[3] Denis J. F. Doyle, M. Eng. Sc. Thesis, University College, Cork, 1987.

[4] P. M. Zeitzoff, et al, IEEE Trans. on Electron Devices, Vol. ED-20, No. 2, April 1985, pp. 489-494.

[5] T. E. Zipperian, et al, IEEE Trans. on Electron Devices, Vol. ED-29 No. 2, February 1982, pp. 341-343.

[6] C. G. Cahill, et al, Proc. of 2nd Int. Conf., Swansea, July 1986, (Pineridge Press, Swansea).

Session P4.1

Posters

Thursday, September 17, 1987

P4.1.2

EMITTER CONCENTRATION AND INTERFACE EFFECTS IN THE BASE-EMITTER CHARACTERISTICS OF AlGaAs/GaAs HETEROJUNCTION BIPOLAR TRANSISTORS

P.Cámara, E. Muñoz, I. Izpura ,J. Lablanca and E.Lapeña
ETSI Telecomunicación, Univ. Politécnica Madrid, C.Universitaria
28040-Madrid , Spain and

M.A.Pate, G. Hill, P. Mistry, J.S. Roberts, and H.Y. Hall
SERC III-V Semic Facility, Dept of Electronic Engineering
Univ of Sheffield,Mappin Street,Sheffield S1 3JD, United Kingdom

Heterojunction bipolar transistors (HBT's) are of great interest for applications in analogue and digital integrated circuits (IC's). Their excellent gain, speed , uniformity and power characteristics are due to the injection properties of the base-emitter (BE) heterojunction (HJ), to the very low base resistance achievable , and to the inherent properties of bipolar devices. Because of its more advanced technology, most of the efforts have been concentrated on AlGaAs/GaAs HBT's. Its technological implementation has made steady progress through the use of graded emitter-base interfaces, and by the intrcduction of superlattices (SL) and offset undoped base layers. Early HBT's made by liquid phase epitaxy (LPE) already showed very high common emitter gains. Because of the improved dimensional control, molecular beam epitaxy (MBE) and metal-organic chemical vapor deposition (MOCVD) technologies should allow improved and more uniform characteristics. However ,the role of technology dependent traps and interface charges may then become a key factor.

Concerning HBT current gain analysis, following the standard concepts of emitter injection efficiency and base transport factor, there have been reports trying to determine the role of emitter concentration, B-E compositional grading , and SL layers, in the B-E characteristics [1,2,3,4]. One striking result was reported by Chand et al. [3], who showed that, by lowering the emitter concentration to a 5×10^{16} cm-3 level, the B-E injection threshold (knee voltage) was decreased from the standard 0.5 V down to 0.1 V. However, all other factors, such as Al grading and the base region doping concentration, introduced practically no significant variations in the forward knee voltage . For such

very low emitter concentration, a high series resistance or current limiting mechanism seems to appear, and no data were reported concerning the achieved transistor characteristics.

In this study a variety of mesa AlGaAs/GaAs HBT's, made by LPE and MOCVD were fabricated . LPE devices have Sn as emitter dopant, while MOCVD devices have Si as donor dopant and Zn as base acceptor. For the emitter region the standard 30% Al composition is used. Base doping level was in the 10^{18} cm-3 range, and emitter free electron concentration was $3x10^{17}$ cm-3. Abrupt, graded and SL, BE interfaces have been considered , and a range of undoped , base offset layer thicknesses, from 0 to 600 Å, have been used.

BE junctions were analysed through I-V , C-V, C vs. frequency and DLTS techniques. One-dimensional device computer simulations were also made . The detection of DX centres in the emitter region was used as a probe to know where the depletion region was being examined, and to monitor hole injection from the base. It is the objective of this paper to show how the various BE interfaces influence their current-voltage characteristics and their deep level structure. It is claimed that the presence of BE interface charge may set a practical limit to the useful minimum emitter concentration, and that very low values in the BE knee voltage can be achieved with emitter concentrations in the 10^{15} - 10^{16} cm-3 range, but the devices do not show transistor effect.

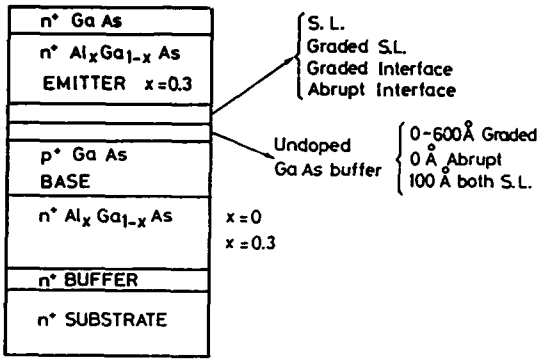


FIGURE 1

The structures being examined are schematized in Figure 1. For different BE interfaces, the base-emitter I-V characteristics are summarized in Figure 2. On a stastistical basis, all the interfaces having an offset layer, tend to increase junction currents at very low voltages, as compared to LPE abrupt or graded interfaces without offset layer. In the injecting range, at similar current densities, junction voltage differences were in the tens of mV range . The introduction of base undoped layers have a more clear effect on the BE I-V characteristics. Figure 3

indicates that in the n≈2 region, at constant Vbe, current increases monotonously with layer thickness, while in the n≈1 regime all the non-zero thickness undoped layer devices have the same behavior.

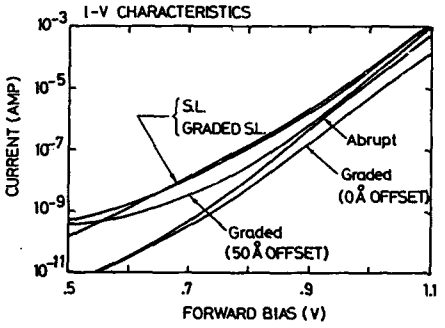


FIGURE 2

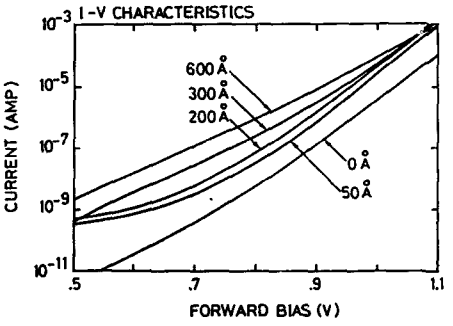


FIGURE 3

The effect of emitter concentration on the BE knee voltage is displayed in Figure 4. At currents below the mA range a significant reduction in the junction voltage is found, a similar result to those reported by Chand et al.[1]. Such region is followed by a high series resistance regime.

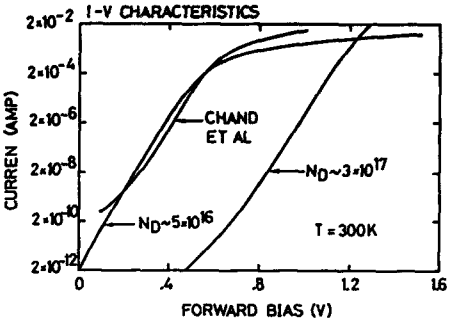


FIGURE 4

Above interfaces were analyzed by DLTS techniques. Because the starting p+-n structures, the emitter regions would display the DX centres linked to the n-type dopant. Due to the DX-centre very large hole capture coefficient [4,5] , monitoring DLTS peak height with forward bias allows to determine the onset of base hole injection into the emitter. A second electron trap was detected,related to the base Zn-doping, its concentration decreasing with the offset undoped layer thickness, and being negligible for offset layers thicker than 300 Å (Figure 5).

Transistors having the various BE interfaces previously analyzed , were evaluated for DC characteristics. Our results indicated that very low emitter concentration devices did not show transistor efect; the minor differences

observed in the BE characteristics, between superlattice and graded interfaces, seem producing neither different DLTS spectra nor transistor characteristics; the undoped offset layer has a more pronounced effect on the I-V characteristics and transistor current gain (Figure 6); best transistor performances were obtained for just zero thickness undoped offset layer and graded interface.

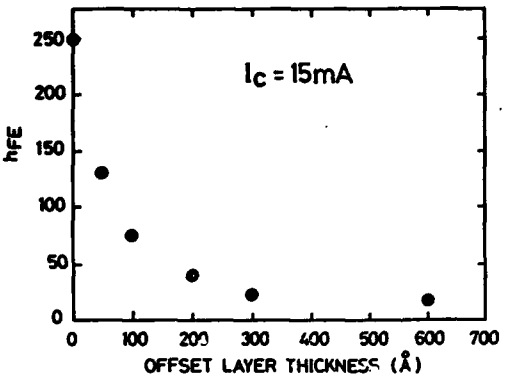


FIGURE 6

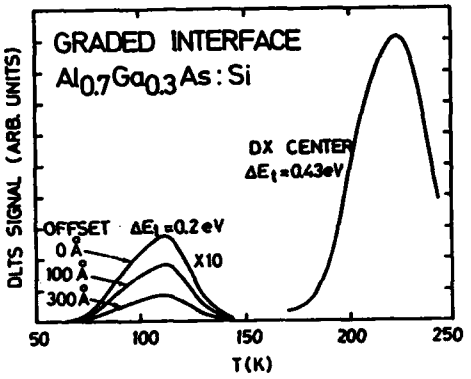


FIGURE 5

REFERENCES

1.-K.Taira,C.Takano,H.Kawai, and M.Arai,Appl.Phys.Lett 49,1278,(1986).

2.-R. Fisher, T. Henderson, J. Klem , N.Chand and H.Morkoc, Solid-St. Electron.29,193,(1986).

3.-N. Chand, R. Fisher, J. Klem and H. Morkoc, J. Vac. Sci. Technol., B4 (2),605,(1986).

4.-S.Tiwari, S.L. Wright, and A.W. Kleinsasser, IEEE ED-34, 185, (1987).

5.-E.Calleja, E. Muñoz, B. Jimenez, A.Gómez,F.García,and F. Kellert, J.Appl. Phys.,59,5295,(1984).

A COMPUTER AIDED INTERPRETATION OF MOBILITY PROFILE MEASUREMENTS IN GALLIUM ARSENIDE FET STRUCTURES

P4.1.3

M. Pillan - F. Vidimari

Telettra - Telefonia Elettronica e Radio s.p.a.
20059 Vimercate (Milan), Italy

Some semiquantitative correlations among electron mobility profile trend, background acceptor and donor concentrations and compensation ratio in gallium arsenide epitaxial FET structures were ascertained from the comparison of several experimental and calculated data.
The results obtained by this analysis are a useful tool in the electrical characterization of non intentionally doped buffer layers where high resistivity and interface (or surface) effects make quite difficult the interpretation of the most used techniques as differential C-V, magnetotransconductance and Van der Pauw measurements.

1 INTRODUCTION

The electrical properties of unintentionally doped gallium arsenide epitaxial layers are generally governed by the presence of background acceptors and donors and by their compensation ratio ($\emptyset = Na/Nd$).

The direct electrical evaluation of these layers is quite critical in the case of interest (very high resistivity) since the presence of a significant surface state density produces a complete free carrier depletion. Furthermore, separate characterization of active and buffer layers does not give any information on the doping profile transistion region.

In order to evaluate the correlations between ionized impurity concentration, compensation ratio and electrical measurements we simulated several different monodimensional active layer-buffer layer structures, numerically reproducing experimental free carrier concentration and mobility profiles.

2 EXPERIMENTAL

The epitaxial structures were grown in a chloride VPE reactor; the buffer layer thickness ranged from 2 to 4 μm; the residual impurity concentration was varied using different GaAs sources. The active layer doping concentration (sulfur) was between 1E17 and 2E17 cm-3.

SIMS analysis were performed on significant samples in order to monitor the doping profile steepness and the impurity concentration.
Residual background impurity concentration resulted to vary from 1E15 to 1E16 cm-3 from sample to sample.

Test patterns including a fat FET, a one micron gate FET and a Van der Pauw cloverleaf structure were manufactured on the wafers under test to perform the electrical characterization.
Free carrier concentration profiles were measured by the differential C-V technique.
Electron mobility profiles were obtained both by C-V and magnetotransductance (1) methods.
From Van der Pauw configuration measurements sheet resistivity and mean mobility values were obtained.

The experimental mobility profiles followed different paths, typically warying from a continuously decreasing trend in samples with highly compensat[illegible]er layer, to a slightly increasing t[illegible]eply decaying one in the most pure structures. Two examples of the above mentioned behaviour are reported in fig.1 and 2 where are also plotted the corresponding free carrier concentration profiles.

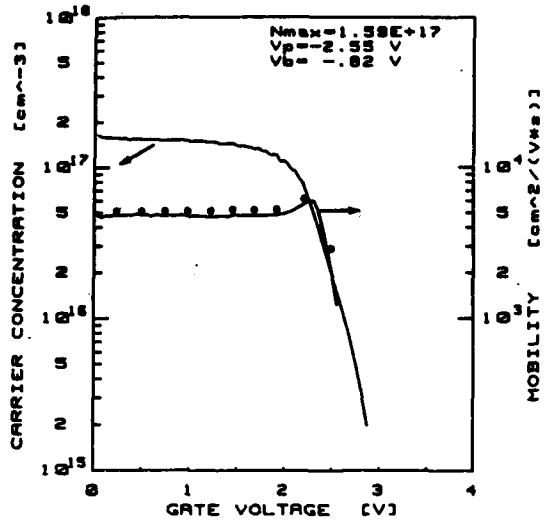


Fig. 1 : Experimental free carrier concentration and mobility profiles, and calculated (●) mobility a low residual impurity concentration sample (Ni IE15 cm^{-3})

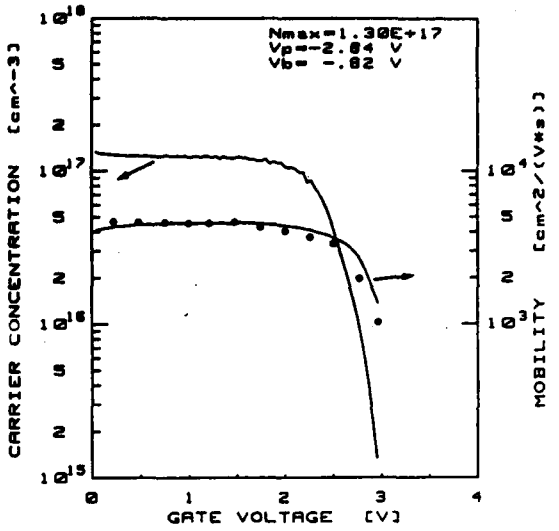


Fig. 2 : Experimental free carrier concentration and mobility profiles, and calculated mobility values (●) for a compensated sample (Ni IE16 cm^{-3}; Na 5E15)

3 COMPUTER SIMULATION

In the computer simulation a typical active-buffer layer structure was examined.

Free carrier concentration profiles were calculated by solving numerically the Poisson operation.

$$\psi'(X) = -q(Nd(x) - Na - \exp(q\psi/kT))/\varepsilon_s$$

Different residual acceptor and donor concentration values, doping level and transition steepnesses were assumed. The Poisson equation solution was performed using a partially modified finite difference method, which reduced the problem to the solution of a linear system equation (2).

The free carrier concentration profiles reported in fig.3 and 4 were calculated assuming a similar doping profile but different residual acceptor and donor concentrations. Significant differences can be observed in the doping profile transition region. The acceptor concentration controls the diffusion of the electrons into the buffer layer, producing a strict carrier confinement for high compensation ratios.

Mobility profiles were numerically obtained by calculating for several boundary conditions, the average electron mobility $\mu(V)$ experienced by the free carriers over the whole electron distribution, as

$$\mu(V) = \int \tilde{\mu}(x)n(x)dx \; / \int n(x)dx$$

Here $\tilde{\mu}(x)$ is the local value of the electron mobility which was deduced from the theoretical data presented by Waluckiewizc (3) as a function of the local ionized impurity concentration and compensation ratio and assuming an effective value

$$\vartheta^*(x) = Na^*/Nd^*$$

in the partially depleted regions where

$$n(x) = Nd^* - Na^*; \; Ni = Nd + Na = Nd^* + Na^*$$

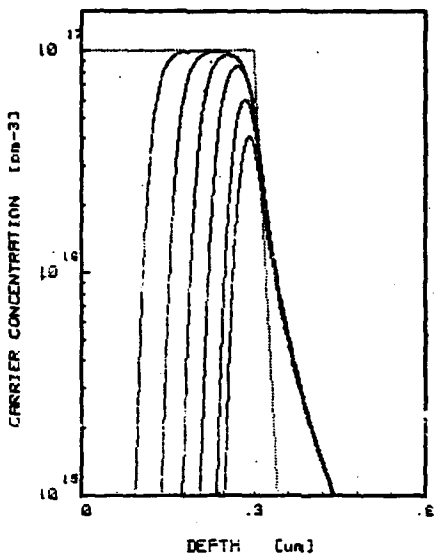


Fig. 3 : Assumed doping profile (·····) and calculated electron distributions for different boundary conditions - Na = 5E14 cm^{-3} ; Ni = 1E15 cm^{-3}

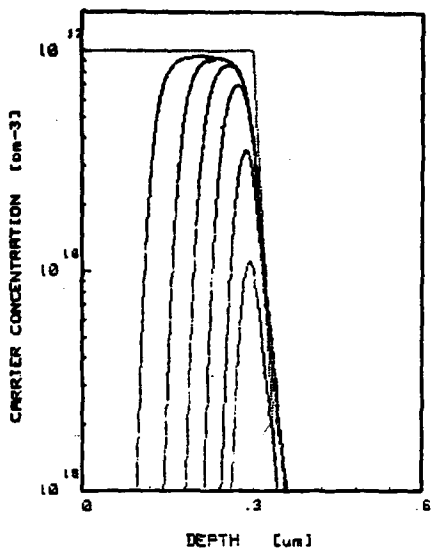


Fig. 4 : Assumed doping profile (·····) and calculated electron distributions for different boundary conditions - Na = 3E15 cm^{-3} ; Ni = 4E15 cm^{-3}

4 DISCUSSION

In order to reproduce the experimental results, the computer simulations were performed assuming experimental doping profiles as measured by SIMS analysis; both residual impurity concentration and compensation ratio were varied over a wide range. For each sample, the compensation ratio in the active layer was deduced from the measured electron mobility in the neutral region. The residual free electron concentration in the undoped layer was assumed to range from 1E13 to 1E14 in coherence with experimental data measured on very thick buffer layers.

The main results of this analysis can be resumed in the following statements.

- The residual acceptor concentrationstrongly affects the fre carrier diffusion at the interface between active and buffer layers.

In very high purity samples (residual Ni=Nd+Na < 1E15 cm-3), regardless of the acceptor concentration exact value, the electrons considerably diffuse from the active layer into the buffer layer where they experience a very high mobility since the residual ionized impurity are superscreened by the exceedingly high number of free carriers.

As the acceptor concentration grows larger than 1E15 cm-3 the electron distribution is progessively more affected by their presence, that produces the confinement of the carriers in the highly doped region. This confinemen results complete if Na > 4E15 cm^{-3}.

- Regardless of the real value Ni, the electron diffusion creates a space charge region at the buffer-active layer interface that justifies the experimental and theoretical fall at pinch off since in these conditions most electrons are in a region where the scattering phenomena are enhanced by the underscreening of the ionized impurities (4).

- In heavily compensated samples, the electrons in the distribution tail undergo a strong underscreening effect since the ionized

impurity concentration is much larger than that of the free carriers. This effect degrades the mean electron mobility and produces the slowly decaying trend reported in fig.2.

The steepness of the mobility profile and its mean value are controlled by residual Ni and by the doping profile trend.

- In very pure samples, as can be observed in fig.3, the percentage of diffused electrons into the buffer layer, becomes more and more significant as the depletion voltage approaches pinch off.

The average electron mobility $\mu(V)$ results therefore a growing function of the bias voltage V as far as the electron concentration does not fall under the background ionized level. Here again the mean mobility value and the amount of its raise are governed by both residual impurity concentration and compensation ratio. The position of the abrupt mobility profile decrease seems to be a function of the doping profile knee shape (exponential, gaussian, etc.)

A further proof of the validity of this model is given in fig.5 where we report carrier concentration and mobility profiles measured on a low background structure by strongly negatively biasing the substrate to produce a backside gating effect on the active layer. In this condition the mobility enhancement disappears because all the carriers result confined in the active layer.

5 CONCLUSIONS

The experimental mobility profiles on gallium arsenide FET structures can be interpreted in terms of buffer layer background impurity concentration and compensation ratio, taking into account non neutrality effects at the interface with the active layer.

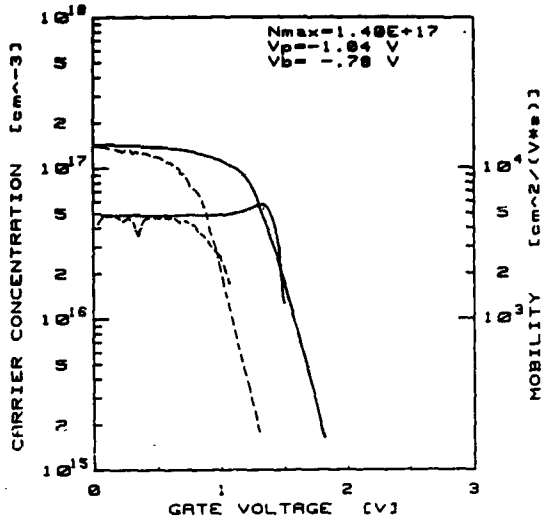


Fig. 5 : Experimental free carrier concentration and mobility profiles measured for 0 V (——) and -30 V (— — —) backside gating bias

Investigations on the influence of the buffer layer characteristics on microwave FET's device performances are in progress, with special attention to pinch-off and beakdown voltages.

REFERENCES

(1) Jay, P.R., and Wallis, R.H., IEEE Electron Dev. Letters EDL/10 (1981) 265

(2) Klopfeinstein, R.W., and Wu, C.P., IEEE Trans. on Electron Dev. ED222/6 (1975) 329

(3) Walukiewicz, W., Lagowski, J. Jastrzebski, L., Liechtensteiger, M., and Gatos, H.C., Journ. Appl. Phys. 48/2 (1979) 899

(4) Pillan, M., and Vidimari, F., in: Kukimoto, H., and Miyazawa, S., Semi-Insulating III-V Materials, Hakone (North Holland, Amsterdam, 1986) 585-590

PLANAR GAAS MIXER DIODES FOR MILLIMETER WAVE MMIC's

J. Selders, A. Colquhoun and K. E. Schmegner[+]

TELEFUNKEN electronic GmbH, Heilbronn, West Germany
[+]AEG, Ulm, West Germany

A new fabrication process for planar GaAs mixer diodes operating in the millimeter wave range of frequencies is presented. This technology, which is fully compatible with MESFET fabrication on the same chip, combines the advantages of selective ion implantation for the n^+ contact regions and epitaxial growth for the active n-layer. With Si implantation n^+-layer sheet resistances of down to 16 Ω/□ have been achieved using a rapid thermal anneal. Schottky diodes with Al and Ti Schottky contacts have been fabricated with different contact areas and configurations. Diode resistances of 8 Ω and junction capacitances of around 20 fF were obtained for a contact area of 1.5 x 6 μm^2. n-factors were typically 1.15 for Ti and 1.3 for Al Schottky contacts. RF performance was tested in a hybrid mixer configuration with an oscillator frequency of 35 GHz. Conversion losses of 6 or 5.5 dB and noise figures of 5 or 4.5 dB have been determined for Al or Ti Schottky diodes, respectively.

1. INTRODUCTION

There is great interest in receivers for millimeter wave frequencies for applications such as phased array radar systems. However, the performance of these receivers strongly depends on the quality of the nonlinear mixer element. GaAs Schottky diodes have already shown excellent RF performance up to the millimeter wave range of frequencies /1/. For integration with other active devices such as GaAs MESFETs and passive components in an MMIC, a rather complicated fabrication process has often to be used, giving a compromise between the different requirements of the active devices. Generally the n^+-n-layer sequence is fabricated either by epitaxial growth of both layers /2/ or by ion implantation /3/. Both processes have severe restrictions. With epitaxial layers, low resistance, thick n^+-layers can be easily combined with a high quality active n-layer. Planar structures can however only be realized by a difficult selective growth. Ion implantation allows a simple means of selectively doping combined with planar surfaces. However, n^+-layers with low sheet resistances and moderate doping levels at the surface are difficult to realize.

We present a novel process which combines the advantages of selective ion implantation with the growth of high quality active layers by an MOCVD process. Based on this technology which is fully compatible with the fabrication of MESFETs and MMICs on the same wafer /4/, planar mixer diodes for millimeter wave applications were realized.

2. DIODE TECHNOLOGY

Starting with undoped s.i. GaAs substrates a selective ion implantation with Si was performed using SiON as a mask (see fig. 1). A multiple energy implantation with maximum energies of 360 keV and total doses of up to 1.4×10^{15} cm^{-2} has been used. Two different annealing processes have been tested. One is a conventional furnace anneal (FA) at elevated temperatures of up to 950 °C for 15 min using a GaAs proximity cap. The other process is a rapid thermal anneal (RTA) at 1050 °C for times of 5 sec in N_2 using SiO_2 or SiON as a protecting cap. After anneal and appropriate surface preparation an n-type layer with a doping concentration of $n = (1.5 - 2) \times 10^{17}$ cm^{-3} was grown directly onto the selectively implanted substrates without any buffer layer /4/. Isolation of the active layer between the diodes was performed with a boron implantation,

using SiO_2 as a mask. 1 µm Al was evaporated and the active contact areas were defined with optical lithography. A well known self-aligned technique was used to accurately control the dimensions of the Schottky contact and the separation between the Schottky contact and the AuGe ohmic contact. 200 nm plasma Si_3N_4 was used to passivate the GaAs surface. A Ti/Pt/Au metallisation formed Schottky contacts for the Ti Schottky diodes and was plated to a thickness of around 2 µm for the air bridges and strip lines.

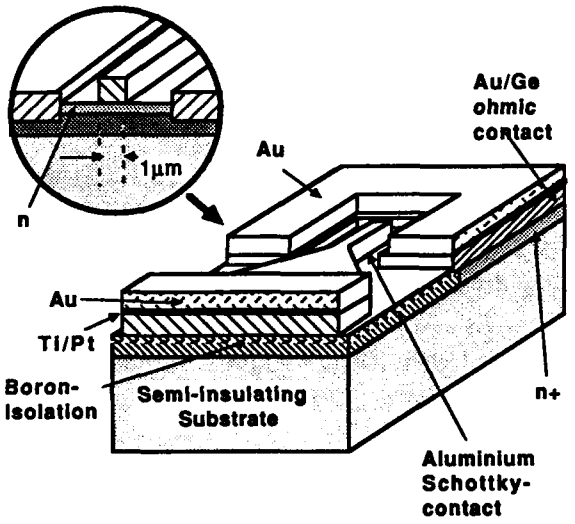


Fig. 1: Schematic view of a FET compatible Schottky diode

3. *DIODE OPTIMISATION*

For mixer diodes operating in the millimeter wave range of frequencies low series resistances and capacitances are required. Therefore special care was taken to optimize the n^+-sheet resistance formed by Si ion implantation. Conventional furnace anneal at temperatures of 850 °C gives sheet resistances for our implantation energies of around 60 Ω/□ (see Tab. 1). A distinct improvement could be obtained by increasing the annealing temperature to 950 °C. A GaAs proximity cap and additional AsH_3 in the annealing atmosphere prevented surface degradation. Minimum sheet resistances of about

process	T/°C	t/sec	cap	lowest R_S/Ω/□
FA	850	900	GaAs	60
	900	900	GaAs	40
	950	900	GaAs	30
RTA	1050	5	SiO_2	22
	1050	5	SiON	16

Tab. 1: Lowest n^+-layer sheet resistance for different annealing conditions

33 Ω/□ with doses of 4.6×10^{14} cm^{-2} could be achieved (see fig. 2). Selective Hall measurements show that a maximum concentration of $n = 2 \times 10^{18}$ cm^{-3} is maintained down to a depth of about 0.5 µm together with electron mobilities of around 1500 cm^2/Vs. A further improvement of the sheet resistance could be obtained by applying rapid thermal annealing using an AG 210 T heatpulse system. The samples were capped with 100 nm SiO_2. Peak temperatures of 1050 °C for 5 sec gave a mini. µm sheet resistance of 22 Ω/□ (fig. 2).

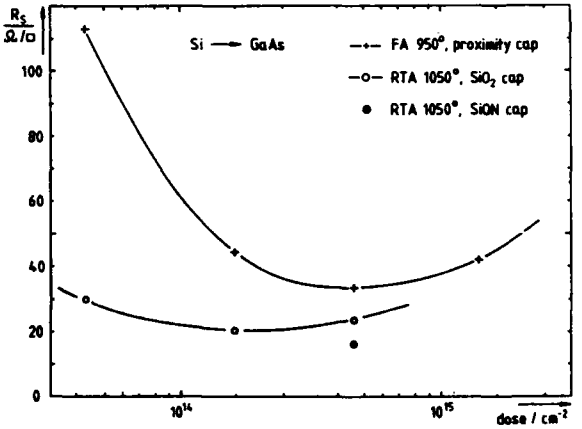


Fig. 2: n^+-layer sheet resistance versus implantation dose

Capping with 100 nm SiON has resulted in a sheet resistance as low as 16 Ω/□ , for a dose of 4.6×10^{14} cm^{-2}. Fig. 3 shows that the decrease in sheet resistance is mainly caused by an increase in maximum carrier concentration above the value of $n = 2 \times 10^{18}$ cm^{-3} which is thought to be the solutibility limit for fur-

nace anneal of Si-implantation in GaAs /5/. With the RTA process, however, maximum concentrations of 5×10^{18} cm^{-3} could be reached. The mobility compares well to the values in case of furnace anneal (fig. 3).

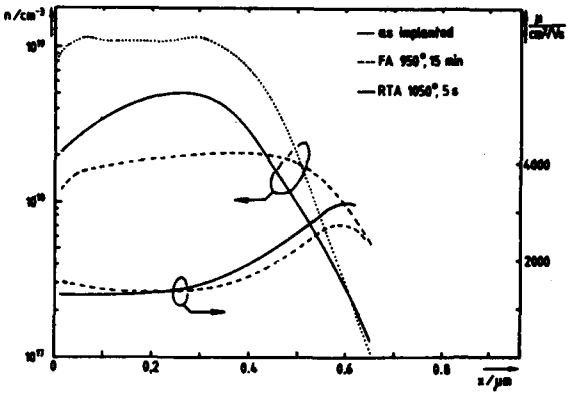


Fig. 3: Concentration profile as obtained after selective Hall measurements

With such sheet resistances a considerable part of the remaining diode series resistance is due to the ohmic contact resistance. Therefore the alloying process has also been optimised giving typical specific contact resistances of around 3×10^{-6} Ωcm^2. Fig. 4 shows the influence of the n^+-sheet resistance and contact resistance on the diode series resistance based on a calculation after Heaton /6/. The figure shows clearly that for our technology diode resistances for contact areas of 1×10 μm^2

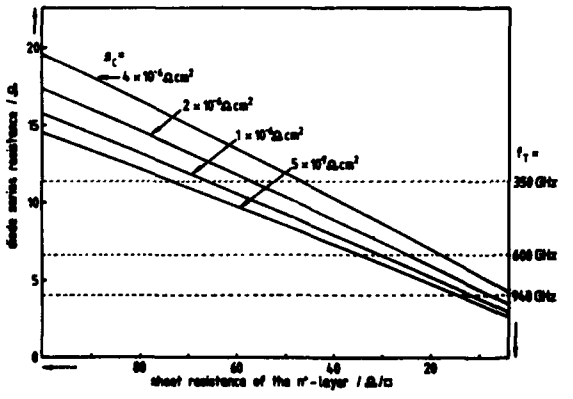


Fig. 4: Diode series resistance versus n^+-layer sheet resistance and specific contact resistance after /6/.

and sheet resistances of typically 25 Ω/□ should result in values of about 7 Ω. This gives f_T-values of about 600 GHz for such diodes making them suitable for operation at frequencies well above 35 GHz.

4. DIODE PROPERTIES

15 different diode configurations such as finger diodes, air bridge diodes with Al or Ti Schottky contacts have been realized in a diode array (see Tab. 2).

Schottky metal	area/ μm^2	structure	mean n-factor	meas R_S /Ω	$C=C_j+C_S$/ fF	f_T/ GHz
Al	1.5·6	finger	1.35	8	30	560
	2.5·6		1.33	8	39	410
	3.5·6		1.39	8	48	320
Al	1.5·8	finger	1.28	7	36	560
	2.5·8		1.30	7	48	405
	3.5·8		1.29	7	60	310
Ti	2·5	airbridge	1.14	8	26	650
	3·5		1.15	8	33	490
	4·5		1.17	8	38	410
Ti	2.7	finger	1.13	11	41	480
	3·7		1.13	11	49	365
	4·7		1.18	12	60	290

Tab. 2: Typical properties of different Schottky diodes

DC measurements were made with the help of an automatic prober so that the properties of each type of diode could be mapped over the wafer. By measuring the I-V-characteristics, the series resistances, n-factors and leakage currents have been determined. Additionally the capacitance of each diode including the strip line capacitance of about 20 fF was measured. Table 2 shows typical (not best) values for each type of diode indicating that the low series resistances expected from fig. 5 could be reached taking the geometrical difference into account.

f_T values for each diode have been calculated by estimating the junction and stray capacitances. Values of more than 600 GHz underline the good properties of these diodes. Al diodes typically show somewhat higher n-factors of 1.25 - 1.3. However, the self-aligned technology in case of Al diodes leads to nearly 100 %

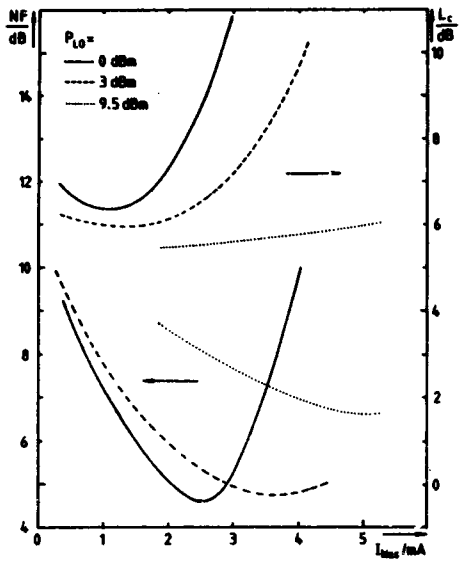


Fig. 5: Noise figure and conversion loss of a Ti Schottky diode versus bias current and LO power at 35 GHz

yield over the waver. The lower n-factor of typically 1.15 for Ti diodes make them even more suited for mixing applications.

The highest f_T values were obtained with air-bridge connections because of lower stray capacitances. However, also Al diodes with finger lengths of 1.5 µm exhibit f_T values of about 600 GHz.

RF performance of such diodes has been tested at 35 GHz in a hybrid mixer configuration. Conversion loss and noise figure were measured at different bias currents and LO powers. Fig. 6 shows a typical result for a Ti Schottky diode. Minimum conversion losses of 5.5 dB and minimum noise figures of 4.5 dB at 2.5 mA bias have been measured. Al diodes show somewhat higher minimum values for the conversion loss of 6 dB and noise figure of 5 dB. These results demonstrate that our FET compatible technology is well suited for MMIC applications at millimeter wave frequencies.

SUMMARY

GaAs Schottky diodes have been fabricated for mixer applications using a FET compatible planar process which combines selective ion implantation and epitaxial growth. Si implantation with maximum energies of 360 keV and RTA resulted in sheet resistances of as low as 16 Ω/□ . In this way Al and Ti Schottky diodes were fabricated with series resistances of about 8 Ω for a contact area of 1.5 x 6 μm^2 and n-factors of typically 1.3 and 1.15 respectively. In a hybrid mixer configuration Ti Schottky diodes exhibited a conversion loss of 5.5 dB and a noise figure of 5.0 dB.

ACKNOWLEDGEMENTS

The authors would like to thank G. Ebert, H. Renz, and R. Stowasser for their help and the German "Bundesverteidigungsministerium" for financial support.

REFERENCES

/1/ J.A. Calviello: "GaAs Schottky Barrier Devices and Components", Microwave Journal, (August 1979), 53

/2/ R.L. von Tuyl: "A monolithic GaAs FET RF signal generation chip" IEEE B SSC Dig. Tech. Papers (San Francisco, 1980), Vol. 23, p. 118

/3/ G.K. Barker, M.H. Badowi, J. Mun, "60 GHz Monolithic GaAs Front-End Circuit for Receiver Applications", Electron. Lett. 20 (8), 1984, 334

/4/ G. Ebert, A. Colquhoun: "High transconductance OGFETs", Proc. of the SPIE Conf. Advance Processing of Semiconductor Devices", Bay Point 1987, to be published

/5/ T.C. Banwell, M. Maenpaa, M.-A. Nicolet, J.L. Tandon: "Saturation of Si Activation at high doping levels in GaAs", J. Phys. Chan. Solids, 44 (6), (1983), 507

/6/ J.L. Heaton: "Analytical model speeds planar-mixer diode design", Microwaves & RF (Sept. 1984) 8 102

LOW-NOISE BULK UNIPOLAR DEVICES IN Si AND GaAs

H.Beneking*, J.-M.Cloos, G.Fernholz, M.Marso, P.Roentgen and L. Vescan

Institute of Semiconductor Electronics, Aachen Technical University, Sommerfeldstr., 5100-Aachen
*Present address: University of Michigan, Department of Electrical Engineering and Computer Science, Ann Arbor

Low-temperature vapour phase epitaxy was applied to fabricate multilayer structures like camel diodes and camel transistors. Using the Low Pressure Vapour Phase Epitaxy (LP-VPE) silicon camel diodes with a conversion loss as low as 6 dB at 12 GHz and a noise figure of 7 dB were realized. GaAs hot electron transistors were fabricated by Organo-Metallic Vapour Phase Epitaxy (OM-VPE).

1. INTRODUCTION

The great improvement in the gas phase epitaxy has lead to the successful fabrication of multilayer structures like camel diodes in both Si [1] and GaAs [2]. Moreover, by planar n-doping V-shaped potentials were built in GaAs, where 2-dim. electrons were proven to exist confined to different subbands.

The camel transistor, which is the first in a new generation of high speed transistors, consists of several deeply buried, very thin, highly doped n and p layers with extremely abrupt interfaces [3]. To realize this device, low temperature processing and no memory effects during the epitaxy are required.

For the fabrication of Si camel diodes the vapour phase epitaxy at reduced pressure (LP-VPE) was used. By this technique a minimum epitaxial temperature of 735°C have been achieved and sharp boron profiles of 13 nm over two orders of magnitude [4]. The particular advantage of this technology, when using chlorosilanes, is that the devices can be isolated by selective growth on Si in windows etched in SiO_2. Si camel diodes with an area as small as $3x3\ \mu m^2$, showing negligible leakage currents at the periphery, were obtained [5]. The diodes had barrier heights in the range 0.46 - 0.94 eV and ideality factors as low as 1.08.

For GaAs we have optimized the organo-metallic vapour phase epitaxy (OM-VPE) in respect to thin and highly doped multilayers by using Mg and Se as dopants and extremely low growth rates [6]. Se allows to adjust donor concentrations greater than $10^{20}\ cm^{-3}$ and sharp transition widths when growing with low growth rate (1.8 nm/min). Using bis-methylcyclopentadienyl-magnesium and a low growth rate, too, of 8 nm/min a precise control over the acceptor concentration is possible. Table 1 shows the main characteristics of the deposition techniques used in this work.

In the following, some recently obtained results related to LP-VPE, respectively OM-VPE grown camel diodes and transistors will be presented.

2. DEVICE FABRICATION

2.1. Si camel diodes

The basic structure of the Si camel diode is shown in fig.1. The diodes were fabricated in 2 inch, (100) wafers 1.5 m Ωcm. First, a 1 μm-thick SiO_2 layer was thermally grown. Then windows with diameters ranging from 3 μm to 400 μm were opened in the oxide using dry etching. After a cleaning procedure the substrates were loaded into the reactor and three layers were grown in one run. These are an undoped n^--layer 0.8-1 μm thick, with an n-type background of $(1\text{-}5)x10^{16}\ cm^{-3}$, a p^+-layer 7-20 nm thick, $(5\text{-}8)x10^{18}\ cm^{-3}$, and a n^+-layer 20 nm thick $3x10^{19}\ cm^{-3}$ doped. At the bottom AuSb was deposited

Table 1. SURVEY OF DEVICE DEPOSITION CHARACTERISTICS

	Si - LPVPE	GaAs - OMVPE
reactor	quarz cold wall	
heating	halogen lamps	
temperature	800°- 823° C	600° C
pressure	0.002 bar	1 bar
gases	$SiCl_2H_2$, H_2	AsH_3, TMG, H_2
doping	in situ	in situ
doping gases	B_2H_6, PH_3	H_2Se, M_2Cp_2Mg
specials	selective growth	two-step growth

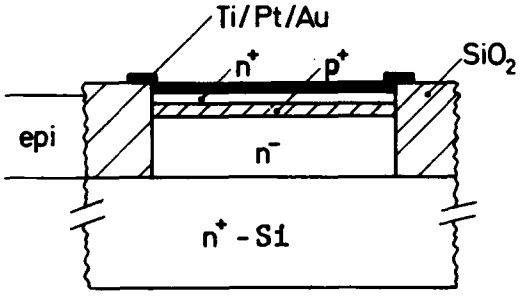


FIGURE 1

Schematic diagram for a selectively grown Si bulk unipolar camel diode.

and annealed at 450°C for 5 min. Then at the top Ti/Pt/Au was deposited and annealed at 300°C for 2 min.

2.2. GaAs hot electron transistor structure

The hot electron transistor structure was grown in two steps in order to be able to contact the thin base. First, the camel collector was grown. It consists of a $1x10^{16}$ cm^{-3} Se doped n^--layer 500 nm thick , a Mg doped p-layer $(5\text{-}20)x10^{17}$ cm^{-3} which is 10-20 nm thick and a Se doped n^+-layer $2x10^{19}$ cm^{-3} , 50 nm thick. This n^+-layer is the base region of the transistor. Then a mesa etching was performed to define the collector area. A CVD deposition at 350°C of 0.1 μm SiO_2 followed. Holes were dry etched into the SiO_2 to define the emitter area. A second epitaxy followed for the camel emitter. The epitaxy sequence was this time $p/n^-/n^+$ with similar doping and thickness values as for the collector diode. During this step GaAs was deposited epitaxially in the holes on GaAs and in polycrystalline form on the SiO_2. After epitaxy, the polycrystalline deposition on the SiO_2 was removed by a chemical etch, and a 1 μm thick SiO layer was deposited by an electron gun. With one photolithographic step the emitter, base and collector contact windows were opened into the SiO. The emitter area was 10x10 μm^2. Then Pt/Si/Au was deposited and contacts were defined by lift-off. A thermal annealing at 350° C was performed to obtain ohmic contacts.

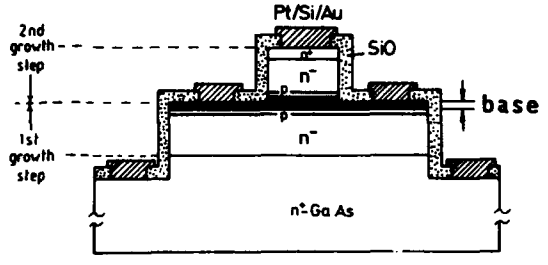


FIGURE 2

Schematic diagram of a hot electron transistor in GaAs

3. DEVICE RESULTS

3.1. Si camel diodes

The carrier profile, recorded by using the Polaron C-V profiler, is shown in fig.3 for a three layer structure grown as follows: an undoped layer was grown at 848°C, then the temperature was lowered to 803°C and 15 nm boron- and 10 nm phosphorus doped Si was deposited. The n^+-layer could not be detected owing the high doping level.

Varying the thickness and/or the concentration of the boron doped region it was possible to fabricate camel diodes with zero volt barrier heights ranging from 0.46 eV to 0.88 eV. Fig.4 shows I-V characteristics of four diodes, grown selectively in 50 μm windows. The diodes have good ideality factors of 1.1-1.2 and low leakage currents. This was verified by measuring diodes with different area deposited in one run and which revealed proportionality between current and area over many orders of magnitude. For noise figure and switching time measurements the diodes were incorporated in micro-strip lines. The diodes were switched from a forward current density of 1000 A/cm^2 to the steady-state reverse current. The measured switching time was below 0.4 ns and no storage phase was detected. This is an indication of the absence of minority carrier effects. For noise measurements the diodes were used as single ended mixers [5]. At a signal frequency of 2 GHz a conversion loss as small as 4.2 dB and a single side band noise figure of 4.8 dB was measured. Diodes with a smaller area, for instance, 8x8 μm^2 revealed a loss of 6 dB and a noise figure of 7 dB up to 12 GHz.

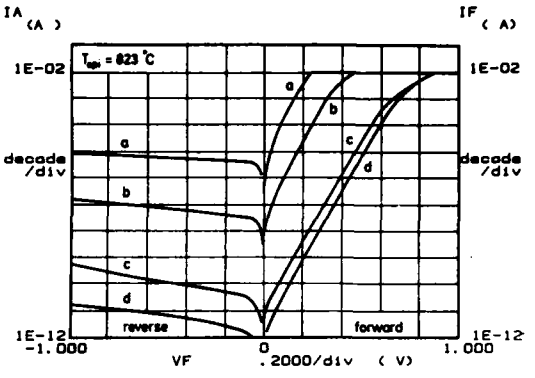


FIGURE 4

I-V characteristics of selectively grown Si bulk unipolar camel diodes. Diode diameter: 50 μm.
a) n= 1.08 and ϕ = 0.46 eV;
b) n= 1.16 and ϕ = 0.59 eV;
c) n= 1.22 and ϕ = 0.80 eV;
d) n= 1.15 and ϕ = 0.88 eV.
ϕ - zero voltage barrier height

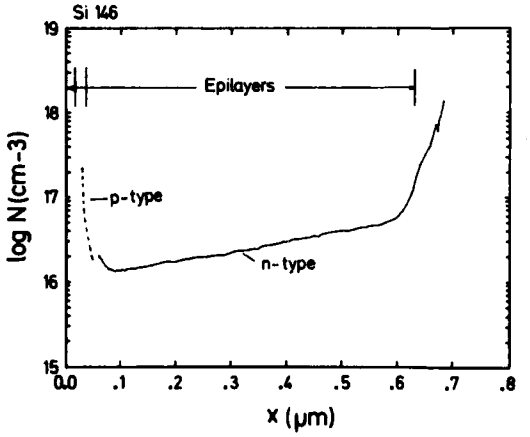


FIGURE 3

Electrochemical C-V profile of a three layer structure in Si ($n^+/p^+/n^-$ on a n^+ substrate). The n^--layer was grown at 848°C and is unintentionally doped. The doped layers were subsequently grown at 803°C.

3.2. GaAs hot electron transistors

To allow the measurement of the transistor doping profile by the C-V electrochemical method, it was necessary to grow thicker n- and p-layers , around 100nm (which is much more than is needed for the camel transistor). Fig.5 shows the layer sequence and the measured do-

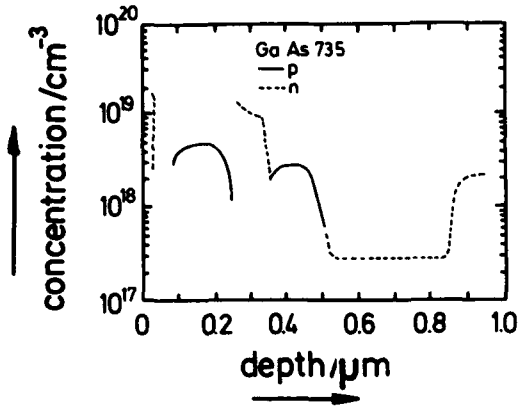


FIGURE 5

Doping profile of a GaAs test structure with thick n^+- (base) and p- (barrier region) regions determined by a Polaron C-V Profiler.



FIGURE 6

GaAs-hot electron transistor characteristic in common emitter configuration measured at room temperature (50 µA/div. vertical, 500 mV/div. horizontal, 500 µA base current steps).

ping profile of a camel transistor test structure. Fig.5 clearly shows that under optimized growth conditions the OM-VPE offers access to this group of hot electron devices.

Fig.6 shows the common emitter characteristic of a camel transistor with a 50 nm base, the maximum current gain is low at 0.16 [7]. This is mainly due to fact that the base is at least three times thicker than the mean free path of electrons. A high increase in the current gain is expected if transistor structures with a much thinner base will be realized. This was demonstrated for MBE-grown GaAs hot electron transistors which showed a gain of 15 [8].

4. CONCLUSIONS

The gas phase epitaxy was succesfully applied for both Si and GaAs to realize low noise Si diodes suitable for high frequency application and GaAs monolithic hot electron transistors.

REFERENCES

[1] Vescan, L. and Beneking, H., Electronics Lett., 22 (1986) 994
[2] Roentgen, P., Fernholz, G. and Beneking, H., in: High Speed Electronics (Springer series in Electronics and Photonics, FRG, 1986) vol.22) pp. 144-147
[3] Shannon, J.M., IEEE J. Solid State and Electron Devices, 3 (1979) 142
[4] Vescan, L. Breuer, U., Werres, Ch. and Beneking, H., in: Proc. of the Int. Symp. on Trends and New Applic. in Thin films, (Societe Francaise du Vide) Strasbourg 1987, pp. 217-221
[5] Werres, C., Vescan, L. and Beneking, H., Electronics Lett., 23 (1987) 613
[6] Roentgen, P., PhD Thesis, Aachen Techn.-Univ., 1986
[7] Beneking H., Roentgen, P. and Vescan, L., "Recent results of OMVPE grown GaAs and LPVPE silicon structures for hot electron devices", Ballistic Electrons for Transistors , Workshop, 1987, Santa Barbara
[8] Woodcock, J.M., Harris, J.J. and Shannon, J.M., Physica B+C (1985) 134

TUNGSTEN SILICIDE RESISTORS FOR GaAs MMICs

D A ALLAN, T K NG and M J GILBERT

Compound Semiconductor Microelectronics Section,
British Telecom Research Laboratories
Ipswich, United Kingdom.

Tungsten silicide resistors in the range 50-300 ohm/square have been deposited on GaAs by rf sputtering and patterned by etching in an SF_6 plasma. The stability of the resistors has been demonstrated by accelerated ageing at elevated temperatures. A change in resistance of less than 0.3% after 1000 hours at 125°C was observed when a Si_3N_4 passivation layer was used to encapsulate the resistors.

1 INTRODUCTION

A variety of different thin film resistor materials have been used in the fabrication of GaAs MMICs, some originating from discrete and hybrid component technology. Ion implanted GaAs resistors are an obvious choice but they suffer from non-ohmic behaviour at fields above $3*10^5$V/m due to velocity saturation and they have a high positive temperature coefficient of resistance. Cermet, nichrome and tantalum nitride thin film resistors have been used as alternatives but have some significant drawbacks in terms of stability, uniformity and patternability etc.

Results are presented on tungsten silicide, a material more commonly used as a refractory metal gate in self-aligned MESFET technology [1], to show that it can overcome many of these disadvantages. Its electrical properties, patterning by plasma etching and stability on lifetesting will be described to illustrate the advantages of this new resistor material.

2 EXPERIMENT

2.1 Deposition and Patterning

The resistor material was deposited by rf diode sputtering of a silicon target with a tungsten sheet overlay which had holes in it to vary the stoichiometry of the silicide film. A composition of W:Si = 1.6:1.0 was chosen to give a sheet resistance of 50-300 ohm/sq. for thicknesses in the range 100-600 nm. The deposition rate was 0.2nm/s at a power of 200W. The resistors were patterned using an SF_6 plasma and a photoresist mask in a parallel plate reactor. The etch rate of the silicide was 4nm/s at a pressure of 266mbar and a power of 50W.

A metal contact layer of Ti-Pt-Au was deposited by electron gun evaporation and defined by photolithographic lift-off.

2.2 Measurements

A 4 point probe was used to measure the sheet resistance of films and the temperature coefficient of resistance (TCR) was determined by heating to 150°C. Patterned resistors were measured using an automatic parametric tester in conjunction with a stepper prober. The patterned resistors were subjected to accelerated lifetesting at 125°C and 350°C for periods up to 1000hrs and changes in resistance noted. Auger electron spectroscopy (AES) was used to analyse the reasons for any changes occurring.

2.3 Ageing and Passivation

Most thin film resistors require curing at 300-400°C in air in order to obtain stable resistance values [2]. In addition, the resistors are sometimes encapsulated in a passivating layer of, for example Si_3N_4 [3], to prevent subsequent drift. The effects of both high temperature treatment and passivation of silicide resistors have been assessed in terms of resistor changes during subsequent accelerated ageing at 125°C and 350°C.

3 RESULTS

3.1 Resistivity

A resistivity of $3*10^{-3}$ohm.cm was obtained with a TCR of -265ppm/°C. A sheet resistance of 300ohms/sq for a film thickness of 100nm was used to fabricate high value resistors up to 9kohm. The uniformity across a 50mm diameter wafer was good (Table 1) with yields close to 100%.

TABLE 1 - RESISTOR VALUES ON A 50mm WAFER (54 LOCATIONS).

NOM. R(ohm)	MEAN R(ohm)	STD.DEV %
50	54.7	3.84
500	508.5	2.97
3000	3154	3.36
9000	9442	3.35

Reproducibility from wafer to wafer was +/- 5% for wafers deposited and patterned at different times - Table 2. The variation in values is due to film thickness differences, photoresist mask dimensional variations and plasma etching rate changes.

TABLE 2 - RESISTOR VALUES FROM WAFER TO WAFER

Wafer No.	R50 ohm	R500 ohm	R3000 ohm	R9000 ohm
1	54.7	508.5	3154	9441
2	51.3	509.5	3161	9417
3	47.3	466.0	2728	8177
4	52.3	491.0	2916	8623

The resistor values in table 2 could be trimmed after contacting and measurement in order to decrease the spread from wafer to wafer. This was achieved by plasma etching as for initial patterning but at a reduced power level. Table 3 illustrates the trimming of a wafer, which was initially too low a value of resistance, until the required values were reached. The results show that there is a limit to the fineness to which the trimming can be performed in that there is no change in the resistor values between 60 and 70s, because of the "induction time" of the plasma process.

TABLE 3 - PLASMA ETCH TRIMMING OF RESISTANCE

Etch t(s)	R50 ohm	R500 ohm	R3000 ohm	R9000 ohm
0	28.4	263	1630	4860
60	47.2	417	2568	7539
70	47.3	418	2569	7542
90	54.6	496	3066	8938

A thicker silicide layer of 600nm was used to fabricate resistors for 50 ohm/sq target values - Table 4.

TABLE 4 - 50 OHM/SQUARE RESISTORS ON A 50mm WAFER (54 LOCATIONS)

NOM. R(ohm)	MEAN R(ohm)	STD.DEV %
8.3	8.74	9.14
83	80.81	5.19
500	479.6	6.16
1500	1467	4.22

3.2 Accelerated Ageing

Experiments to establish whether a stabilisation cure in air was necessary as with other thin film resistors showed a decrease in resistivity of 0.3% after 1 hour at 350°C. A corresponding decrease in TCR from -265 to -150 ppm/°C was observed. There were no apparent differences between heating in air or nitrogen. Longer term lifetests at 125°C and 350°C showed some differences between samples which had been pre-aged and those which had not. Differences were also observed between 50 and 300 ohm/sq. resistors and those which had a surface passivating layer.

Figure 1 shows the percentage change in resistance as a function of time for 300 ohm/sq. resistors. As may have been expected the wafer aged at 350°C changed the most whereas the one lifetested at 125°C after a "burn in" at 350°C for 1 hour showed least change in resistance during ageing.

Analysis of the resistors after the end of lifetesting by AES showed that the oxygen content of the films varied dramatically. After 100 hours at 350°C the oxygen content was 20% at the surface with a steep decrease

as the GaAs interface is approached. This compares with a uniform concentration throughout the layer of 5% for an untreated wafer - Figure 2. A trend, dependent on the sheet resistance of the silicide, was also observed - Figure 3. The change in resistance with time became negative rather than positive as the sheet resistance decreased; ie as the thickness of the film increased.

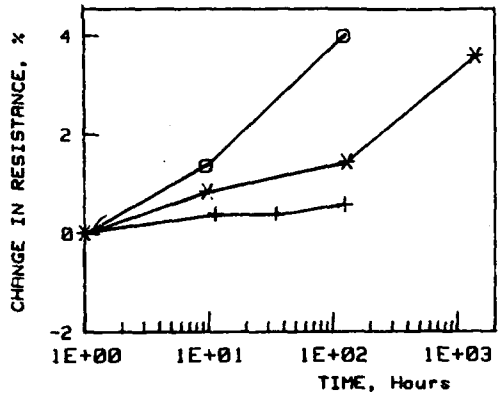


FIGURE 1

Change in resistance of 300 ohm/sq. layers as a function ageing at elevated temperatures:-

* - 125°C

+ - 125°C after pre-age at 350°C for 30 mins

o - 350°C

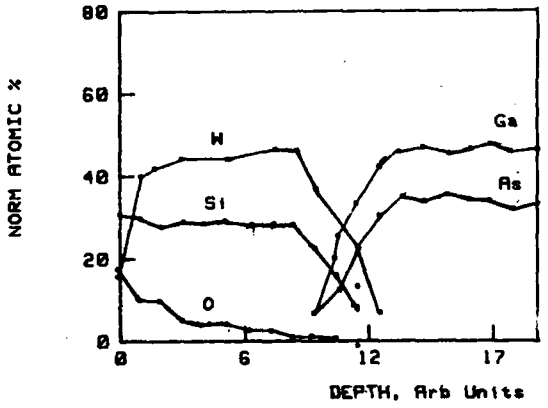


FIGURE 2

AES profile of WSi on GaAs after ageing at 350°C for 100 hours showing oxygen ingress.

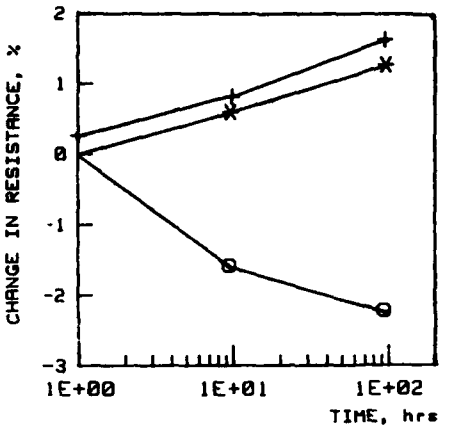


FIGURE 3

Change in resistance as a function of ageing at 125°C for different sheet resistance layers:

o - 50 ohm/sq.
* - 160 ohm/sq.
+ - 300 ohm/sq.

Accelerated ageing of the 50 ohm/sq resistors showed that again the wafer that was not preaged at 350°C showed the largest change with time - Figure 4. The pre-aged wafer did not show such a large change but the minimum variation (less than 0.2%) was achieved when a passivating layer of Si_3N_4 was used on top of the resistors. There was no difference due to pre-ageing when the wafers were passivated since the plasma enhanced deposition of the Si_3N_4 was carried out at 325°C. The wafer with a passivation layer changed by only 0.8% after 100 hours at 350°C.

4 DISCUSSION

The results show that tungsten silicide resistors can be used as alternative to GaAs bulk resistors or cermets for 300 ohm/sq design rules. They are also viable alternatives to NiCr or TaN for 50 ohm/sq applications. They do not have the high positive TCR of GaAs (1500-3200ppm/°C) which can combine with that of active components to give large parameter changes with temperature. Cermet films have been used extensively [4] but in common with other workers we have found problems with stability, reproducibility, and

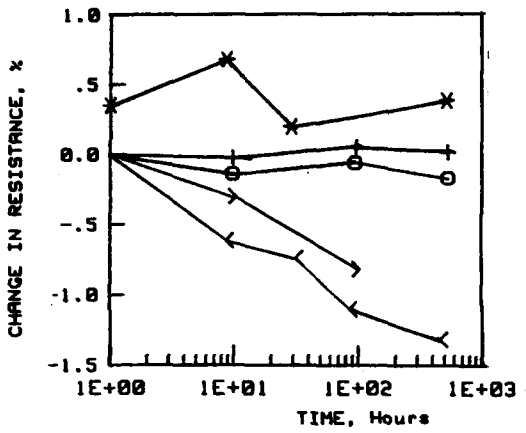


FIGURE 4

Change in resistance of 50 ohm/sq. layers after accelerated ageing:-

< - 125°C
* - 125°C after pre-age at 350°C for 30 min.
+ - 125°C passivated.
o - 125°C passivated and pre-aged.
> - 350°C passivated.

also pattern definition since they are not etched by any of the wet chemicals or plasma processes used in GaAs fabrication technology [5].

NiCr and TaN [2] do not suffer from such great disadvantages. However, WSi is easier to deposit reproducibly compared with TaN, which requires precise control over a small partial pressure of nitrogen for a particular resistivity. WSi is also more compatible with dry etching technology, for precise pattern definition and resistor trimming, than NiCr which cannot be etched in fluorine based plasmas.

The different trends in resistance variation (negative or positive change for 50 and 300 ohm/sq material respectively) with accelerated ageing are explicable in the light of the AES results. The thinner 300 ohm/sq resistors are more affected by the oxygen diffusion into the silicide. The oxygen probably combines with the silicon to form Si-O bonds which causes the resistance to increase. Previous work has shown that WSi deposited by sputtering shows a decrease in resistivity on treatment at elevated temperatures for short periods (30mins) due to crystallisation and grain growth. This is in agreement with results on pre-ageing the resistors and the thicker 50 ohm/sq layers, although a major decrease in resistivity would not be expected below the crystallisation temperature - 650°C.

5 CONCLUSIONS

Tungsten silicide, which is used as a high temperature Schottky gate material in GaAs MESFETs, has been shown to be a stable resistor material for use in GaAs MMICs. It can be reproducibly deposited by rf diode sputtering and is easily patterned by plasma etching. Its stability at high temperatures produces only small drift in resistance during accelerated ageing especially if a passivation layer of Si_3N_4 is used to encapsulate the resistors. It is compatible with GaAs IC fabrication technology and is superior to alternative thin film or GaAs bulk resistors in many respects.

ACKNOWLEDGEMENTS

We would like to thank our colleagues in the Compound Semiconductor Microelectronics Section for their help and advice. In particular we are grateful to Mrs S A Morrice for device photolithography, Dr G Thomas for the AES work and Mr P J O'Sullivan for his guidance of the work. Acknowledgement is made to the Director of Research for permission to publish this work.

REFERENCES

[1] Yokoyama, N., Ohnishi, T., Onodera, H., Shinok, T., Shibatomi, A. and Ishikawa, H. Proc. ISSC83 (IEEE), 1983, pp44-45.

[2] Pengelly, R.S. Microwave Field-Effect Transistors - Theory, Design and Applications (Wiley, New York 1982)

[3] Fraser, A. and Ogbonnah, D., Proc. GaAs IC symposium 1985 (IEEE), 1985, pp161-164

[4] Maissel, L., Thin Film Resistors, in: Maissel, L. and Glang, R. (eds.), Handbook of Thin Film Technology (McGraw-Hill, New York, 1970) pp18-18 to 18-22.

[5] Allan, D. British Telecom Internal Report Unpublished.

MESFETs ON N-GaInAs WITH BARRIER ENHANCED SCHOTTKY GATES

G.Fernholz, W.Lange, R.Westphalen, P.Balk, H.Beneking*

Institute of Semiconductor Electronics, Aachen ,Technical University, FRG

Enhancement of Schottky barrier height on n-type GaInAs has been achieved using Be implantation at low energies.These diodes have been used as Schottky-gates for the fabrication of n-channel MESFETs on GaInAs.

1. INTRODUCTION

GaInAs, lattice matched to InP, is an important material for optoelectronic devices. Due to its excellent electronic properties (high mobility and electron velocity) JFETs with high transconductance and high maximum frequency of oscillation (MAG=1), comparable to MODFET performance, have been achieved /1,2/ . Moreover, the absorbtion edge at 1.67 μm wavelength makes GaInAs an ideal candidate for optoelectronic integration.

It has been proposed that enhancement of the Schottky barrier height may be obtained by using a thin counterdoped fully depleted top layer /3/ . This method has been successfully applied to n-type GaInAs /4-6/ .

In this paper we discuss the fabrication of barrier enhanced Schottky- contacts on n-type GaInAs using Be implantation and rapid thermal annealing (RTA). Such barriers are applicable to the fabrication of MESFETs and can also be used as fast photodetectors /7/.

2.EXPERIMENTAL

Be implantation into n-type GaInAs was performed with energies of 2 to 6keV and doses of $1x10^{14}$ to $7x10^{14}cm^{-2}$. The S-doped GaInAs layers ($1x10^{16}$to $1x10^{17}cm^{-3}$) were 0.2 to 0.5 μm thick and deposited by VPE on semiinsulating InP. RTA was performed in a lamp heated furnace with 13s rise time to the peak temperature of 800°C where it was kept for 1s. As cap material SiO_2 was used, deposited by different methods (evaporation, sputtering or CVD).

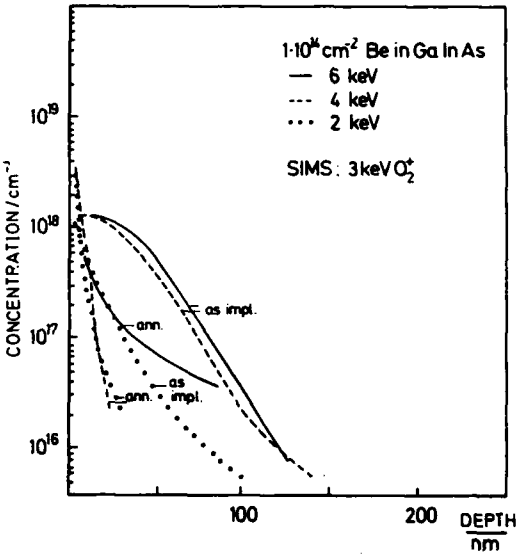


Fig.1: Atomic distribution of 4keV implantation of Be into n-type GaInAs measured by SIMS using 3keV O_2^+ ions

*present Adress: Goebel Visiting Professor, University of Michigan, Dept. of Electrical Eng. and Computer Sci., Ann Arbor, Michigan, USA

Material characterisation was performed by SIMS and PLL measurements. Fig.1 shows the typical atomic distribution of Be, implanted into GaInAs at low energies, where the annealed profiles indicate a strong diffusion towards the surface. A pronounced dependence of the annealed profile on the method used for the SiO_2 deposition could not be detected in our SIMS study. PLL measurements at 4K showed good crystalline quality and an activation of 100%. These results could be confirmed by the excellent agreement between calculated and measured diode parameters (table 1).

Diodes and transistors were fabricated and characterized. For structuring either a mesa ething technique (Fig.2a) or a local implantation of the p^+ region (Fig.2b) was performed. Active device areas were defined by wet etching of the GaInAs down to the InP, whereas the windows opened in the SiO_2 were etched by combining RIE and a wet etching technique. The channel and gate lengths are 3μm and 1.5μm, respectively, at 200μm width. Barrier heights were determined from I-V, I(T)-V and C-V measurements.

3. RESULTS

A barrier height enhancement of 0.42 eV was achieved resulting in leakage currents as low as $2x10^{-3}A/cm^2$ at -9V. This corresponds to a gate leakage of 0.1μA (1μm gate length). An I-V characteristic of a diode is presented in Fig.3.

Table 1: Measured and calculated barrier height enhancement on n-type GaInAs for samples implanted at an energy of 4 keV with different doses.

E keV	Dose cm^{-2}	$\Delta\phi_B$/eV (from J_s)	$\Delta\phi_B$/eV (calc)
4	$1x10^{14}$	0.28	0.22
4	$2x10^{14}$	0.304	0.3
4	$3x10^{14}$	0.33	0.35
4	$5x10^{14}$	0.42	0.41

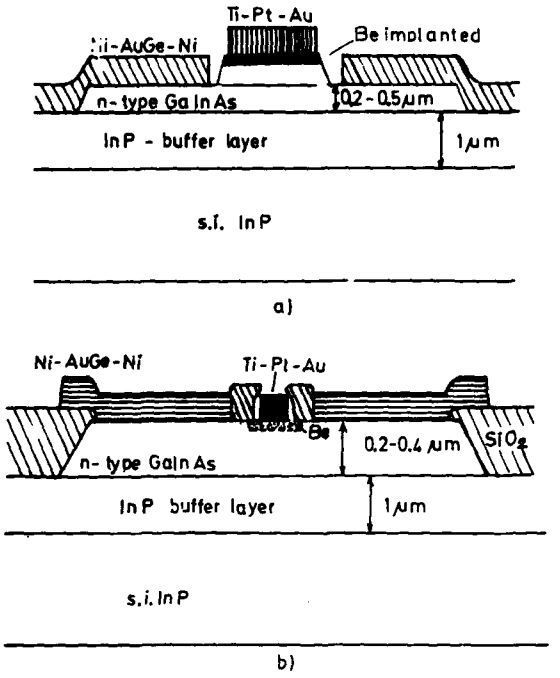


Fig.2: Cross sections of transistor structures a) using an etched and b) using a local implanted gate region

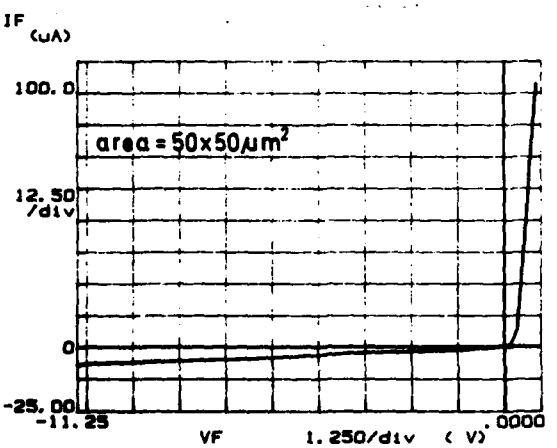


Fig.3: I-V characteristic of a barrier enhanced Schottky-diode on GaInAs

Test transistor channels were fabricated to check the saturation current. From Hall (μ_n=7000 cm^2/Vs) and BPO-profiler measurements and therefore known impurity distribution in the channel, a saturation velocity of $2x10^7$cm/s was calculated. This is a promising value for achieving high transconductances.

The first prepared transistors on 0.2μm thick layers were normally-off at 2V source to drain bias. This resulted in a rather low transconductance of 250μS/mm. Fabrication of devices on thicker active layers is in progress.

4. CONCLUSIONS

Shallow Be implantation with excellent annealing characteristics even at high doses can be used to fabricate barrier enhanced Schottky - diodes and quasi MESFETs on n-type GaInAs. Due to the high electron velocity in the channel achievement of high transconductance transistors may be expected.

ACKNOWLEDGEMENT

The authors are indepted to U.Weimer for metal evaporation, D.Dunkmann for ion implantation, U.Breuer for SIMS measuremets, K. A.Wolter for PLL measurements and A.M.Krings for performing RIE. This work was supported by the German Research Foundation DFG.

REFERENCES

/1/ Schmitt R., Heime K., Electronics Letters 21 (1985) 449

/2/ Raulin J.Y. et al., Appl. Phys. Lett. 50 (1987) 535

/3/ Shannon J.M., Appl. Phys. Lett. 75 (1974) 75

/4/ Chen A.J., Cho A.J., Appl. Phys. Lett. 40 (1982) 401

/5/ Kim H.K. et al., to be published in IEEE Electron Device Lett.

/6/ Fernholz G. et al., to be published in Electronic Letters July 1987

/7/ Emeis N. et al., Electronics Letters 21 (1985) 180

CONDUCTION MECHANISMS ANALYSIS AND SIMULATION OF 1.3 μm LASER DIODES

S. Mottet, A. Changenet*, J.E. Viallet
E. Dudda**, A. Accard**, R. Blondeau***, M. Krakowski***

CNET LANNION, Route de Trégastel - 22301 Lannion - FRANCE
* CNET PARIS, Av. de la République - 92131 Issy les Moulineaux - FRANCE
** CGE-LdM, Route de Nozay - 91140 Marcoussis - FRANCE
*** THOMSON CSF-LCR, Domaine de Corbeville - 91401 Orsay - FRANCE

Numerical simulations and analytical models derived on lasers and test structures enable to determine the parameters which allow to fit experimental I(V) measurements, to pinpoint the semiconductor interface responsible for the excess current and to determine the evolution of the performances under aging.

1. INTRODUCTION

The critical topics that are encountered today in the lasers, for optical fiber telecommunications at 1.3 μm, are increasing efficiency, lowering threshold current and *ensuring the reliability* of the device under operation. Whatever the double heterostructure laser, the *main* difficulties are to obtain reproducible III-V epitaxial layers, to decrease the *leakage* current and its drift and to control the evolution of the device under long term operation. For these reasons it is a necessity to understand which from the epitaxial junctions are responsible for the undesired current and for the degradation of the performance under operation. The main idea is to study separately the active layer diodes and the lateral blocking layer diodes.

2. LASERS AND EXPERIMENTAL STRUCTURES

Two type of 1.3 μm lasers have been studied:
-Type I are BH lasers from Thomson CSF realized by MOCVD epitaxial technique.
-Type II are DCPBH lasers from CGE using the LPE technique for the layer growth.

The test structures which have been characterized have been performed using both techniques, MOCVD and LPE, and are issued from *the same fabrication lines* as the lasers. They are all composed of uniformed layers and contacts on the whole surface. They are cutted in samples of about 300×300 μm^2 surface. There are two types of structure:
-The layers of the A type (IA for MOCVD and IIA for LPE) are those of the laser cross section through the quaternary active layer.
-The layers of the B type (IB and IIB) are the same as type A *without* the quaternary layer.

The unidimensional conduction behavior in these structures of well known surfaces allows the interpretation of the measured electrical characteristics. The A type structures are representative of the only active region of the lasers without confining and blocking layers effects. Both forward and reverse bias characteristics of the B type structures have been measured in order to characterize n p InP junctions which compose the blocking layers in both type of lasers.

3. NUMERICAL SIMULATION

The simulation of the electrical behaviour of the laser is performed by solving the general equation set of the static drift diffusion model to describe heterojunctions under Fermi-Dirac statistics [1]:

$$\begin{cases} \operatorname{div}\, \epsilon \overrightarrow{\operatorname{grad}}\, \varphi = q\,(n\text{-}p\text{-}dop) \\ -\frac{1}{q} \operatorname{div} \vec{J}_n = -U \\ \frac{1}{q} \operatorname{div} \vec{J}_p = -U \end{cases} \quad \text{with} \quad \begin{cases} \vec{J}_n = -qn\mu_n \overrightarrow{\operatorname{grad}}\, \varphi_n \\ \vec{J}_p = -qp\mu_p \overrightarrow{\operatorname{grad}}\, \varphi_p \end{cases}$$

with φ the electrostatic potential and φ_n, φ_p the electron and hole electrochemical potentials.

The carrier densities n, p are expressed using Fermi-Dirac statistics in the parabolic band assumption:

$$n = N_C \mathcal{F}_{1/2}\left[\frac{q\varphi+\chi-q\varphi_n}{kT}\right];\quad p = N_V \mathcal{F}_{1/2}\left[\frac{-q\varphi-\chi-E_g+q\varphi_p}{kT}\right]$$

with χ the electronic affinity and E_g the band gap. The generation-recombination term U takes into account the three main phenomena involved up to the threshold current; $U = U_T + U_{sp} + U_A$:

- *The deep center thermal recombination:* the assumption is made that the whole thermal recombination can be described considering, in each layer, one dominant deep level center. The energy level of this mean deep center is usually considered to be located at midgap. In fact there is no reason for this, and for the active quaternary layer we found that the energy level of the main deep center was above midgap. Under Fermi-Dirac statistics, the formulation is [1]:

$$U_T = \frac{np - n_i p_i}{\tau_p(n+n_i)+\tau_n(p+p_i)} \quad \text{with} \begin{cases} n_i = n \exp\left[\frac{E_T+q\varphi_n}{kT}\right] \\ p_i = p \exp\left[\frac{-E_T-q\varphi_p}{kT}\right] \end{cases}$$

E_T is the energy level of the center, τ_n, τ_p the extrinsic lifetimes of the carriers which depend on the center density.

- *The band to band spontaneous optical emission*: This term is expressed so as to be valid at thermal equilibrium under Fermi Dirac statistics:

$$U_{sp} = B_0\ (np - n_i p_i) \quad \text{with } n_i p_i = np \exp\left[\frac{q\varphi_n - q\varphi_p}{kT}\right]$$

- *The Auger recombination:* this term uses a mean value for hole and electron processes as n = p when Auger recombination occurs.

$$U_A = C_A(n + p)(np - n_i p_i);\ n_i p_i = np \exp\left[\frac{q\varphi_n - q\varphi_p}{kT}\right]$$

All the physicals parameters used in this equation set are temperature dependant so that the simulation can be performed as a function of the device temperature T.

Figure 1 shows the simulated 1.3 µm laser band diagram for .8 V forward bias voltage. It corresponds to high injection level before threshold current. This figure shows that, due to the high carrier densities in the quaternary, Fermi-Dirac statistics have to be used. As a consequence the current voltage characteristic *cannot be describe* by the approximation $I(V) = I_0(\exp(qV/\eta kT)-1)$ which supposes the use of the Boltzmann statistics (whatever the ideality factor η).

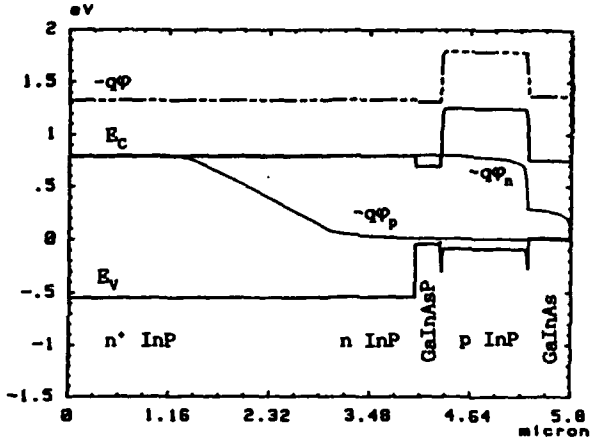


FIGURE 1
Band diagram of a .8 V biased laser.

4. THE PHYSICAL PARAMETERS

Most of the physical parameters such as effective masses and forbidden bandgap energies have been found without ambiguity in the litterature. More attention was taken concerning B_0 and C_A. The 1.3 µm spontaneous band to band emission constant is $B_0 = 10^{-10}\text{cm}^3\text{s}^{-1}(T/300)^{1.5}$. Some discrepancy on the Auger coefficients appears in the litterature. Among the possible values, $C_A = 3.6\ 10^{-27}\exp(-1280/T)\ \text{cm}^6\text{s}^{-1}$ has been chosen ($C_A = 5\ 10^{-29}\text{cm}^6\text{s}^{-1}$ at room temperature). Agreement of these values have been obtained by comparison with I(V) measured characteristics of the A type structures, as a function of the sample temperature. These physical parameters *are fixed for all the computations*.

5. SIMPLIFIED MODEL

From the numerical simulation it has been found that the current due to recombination terms, outside the quaternary active layer, were always negligeable in this type of struc-

ture. So that the current can be directly described by the integral of the recombination terms in the quaternary layer. Carrier densities increase rapidly in the quaternary when increasing the bias voltage and, when $n > 10^{15}cm^{-3}$, the assumption $n = p$, over the whole layer volume v_{quater}, can be made. From these two assumptions an $I(\nu), V(\nu)$ formulation can be derived where the parameter ν is the carrier density into the active layer.

$$I(\nu) = \nu^2 \left(\frac{1}{\tau_0 (2\nu + n_1)} + B_0 + 2C_A \nu \right) . v_{quater}$$

let $\bar{\mathcal{F}}_{1/2}$ be the inverse function of $\mathcal{F}_{1/2}$ which can be numerically computed:

$$V(\nu) = \frac{kT}{q} \left(E_g + \bar{\mathcal{F}}_{1/2}\left[\frac{\nu}{N_C}\right] + \bar{\mathcal{F}}_{1/2}\left[\frac{\nu}{N_V}\right] \right) + R_s I(\nu)$$

τ_0 is the carrier lifetime in the quaternary and n_1 is representative of the energy level of the main recombinaison center . All the physical parameters used in this formulation are those which are used in the numerical simulation and are temperature dependant. The main difference between the two models lies in the serial resistance R_s. The resistivity of the semiconductor material is implicitly taken into account in the numerical simulation (p InP layer), whereas it has to be added in the simplified model. In both cases the contact resistance has to be included. The *only three adjustable parameters* are τ_0 and n_1, which concern the recombination center, and R_s.

6. DISCUSSION

The experimental I(V) characteristic of a laser is plotted in figure 2. This figure shows the result given by the model and the contribution of each of three recombination terms to the total current. It can be seen that each of the terms participates to a different part of the characteristic. The low current range principally depends on the thermal recombination term. The two parameters of this term are τ_0, the lifetime of the carriers, and n_1 which directly depends on the energy level of the main recombination center involved. Changing n_1 changes the slope and the shape of the curve. The lifetime acts on the amplitude of the curve, so that

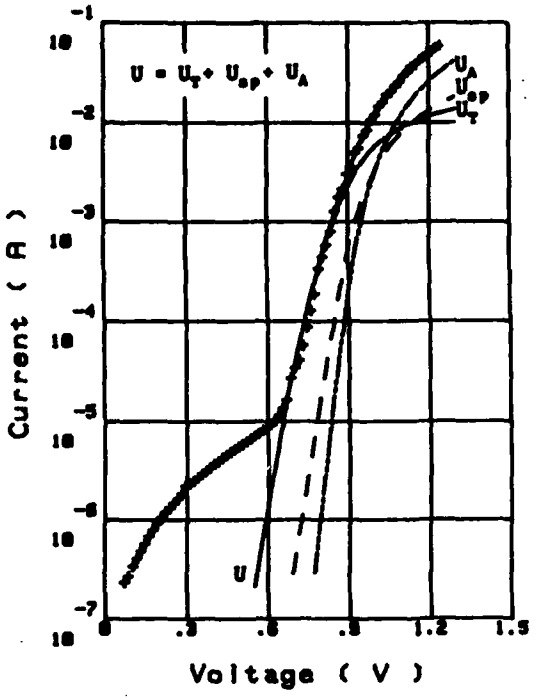


FIGURE 2
I(V) characteristic from experiment and model: contribution of the three recombination terms.

these two parameters are independant and can be deduced by fitting the experimental curves (when fixed they have to fit at any temperature).

Figures 3 and 4 show comparison between experiment and model for the two type of lasers at two temperatures.

The obtained results are very close for both types of laser. The lifetime ranges between some $10^{-8}s$ and $10^{-9}s$, depending on the concentration of the recombination centers.The n_1 value ranges between 10^{16} and $10^{17}cm^{-3}$. This seems that the main recombinaison center energy level is rather close either to conduction or to valence band and seems to be the independant of the epitaxial technique.

Figures 3 and 4 also show that the model is not able to describe the very low currents. This current excess (by regard to the expected current given by the model) was first supposed to take its origin in the lateral blocking layers. For BH lasers, the IB structures give direct caracterisation of the current flowing in the lateral blocking layers. The measured I(V)

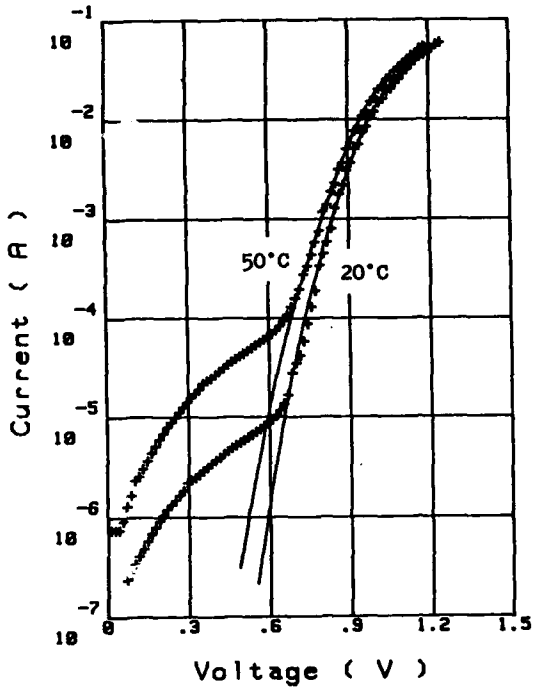


FIGURE 3
Comparison between experiment and model for a type I BH laser at 20°C and 50°C

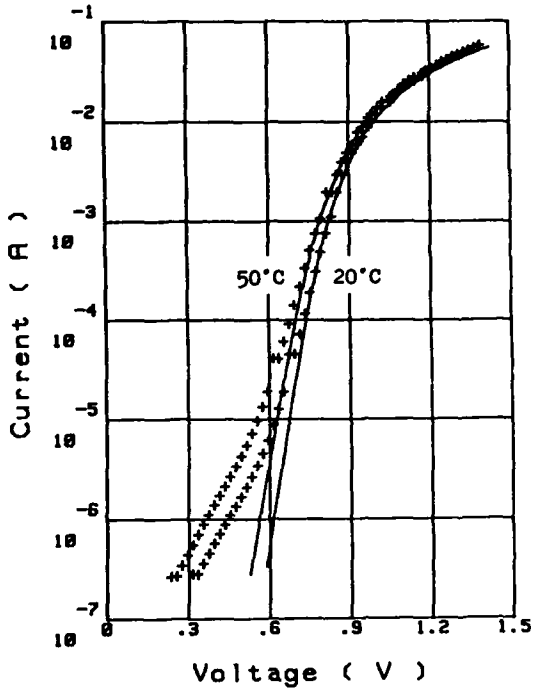


FIGURE 4
Comparison between experiment and model for a type II DCPBH laser at 20°C and 50°C

characteristics show currents some order of magnitude higher than those expected in p n InP homojunctions. Moreover the slope of the curves cannot be describe by the drift diffusion model. The study performed as a function of the temperature led us to think that the current is due to tunnel effect through deep center located at the interface between n InP and the regrowth p InP epitaxial layer. This behaviour is more generally found in abrupt heterostructure than in homojunctions. The fact is that both epitaxial techniques exibit the same behaviour and that it may be conclude to the very high density of defects at the regrowth interface. Nevertheless the magnitude of the current is small enough to be neglected with regard to the current density flowing in the quaternary active layer. Our conclusion is that the excess current does not come from the lateral blocking layers, neither from the quaternary diodes, but takes its origin at the interfaces between quaternary layer and lateral layers.

The difference between measurement and model gives the excess of current and *its evolution* when aging. Moreover the fitted model gives the evolutions of the lifetime in the quaternary and of the serial resistance: typical variation deduced after 20 hours aging at 70°C and 150 mA current are 1.4 10^{-8} to 4 10^{-9}s for τ_0 and 5.5 to 15 Ω for R_s.

7. CONCLUSION

For laser reliability studies, the proposed models permit to identify and separate the main contributions to the current characteristic of the lasers and their evolutions when aging.

REFERENCES

[1] Viallet, J.E. and Mottet, S., Heterojunction under Fermi-Dirac statistics, in: NASECODE IV proceedings (Boole, Dublin, 1985) pp. 530-535

[2] Tsang, W.T., Semiconductors and Semimetals Vol 22, part C (Academic Press, Orlando, 1985)

FAILURE ANALYSIS OF GaAlAs EMITTERS[§]

G. CONTE, F. FANTINI*, F. MAGISTRALI

Telettra S.p.A., Quality and Reliability Dept. - 20059 Vimercate - Italy

M. VANZI

Telettra S.p.A., Quality and Reliability Dept.-Via Capo di Lucca 31-40126 Bologna - Italy

The techniques used in characterizing the GaAs based lasers and LED's are described and failure analysis results are reported for devices coming from incoming inspection, qualification, equipment production and field application. A detailed classification of failure modes and the correlation with failure mechanisms is shown by dividing among external overstresses, package and die-related problems.

1. INTRODUCTION

The reliability of emitting devices is still considered the major limiting factor for the spread of optical communication systems, so that redudancy is used where very strict reliability targets, as in submarine cables, are imposed [1].

Although GaAs based emitters have been used for more than ten years in optical systems, very few reliability data are available, also because they are usually considered company confidential, and the few data published in literature [2-3], are quite unsatisfactory. Moreover, when they arise from really operating equipment, they do not contain any data about the causes of failure [3]. On the contrary, most of the available data come from the accelerated tests [4] so that the correlation between the stress tests and field reliability has not been demonstrated.

In this paper we report the results of failure analyses performed on Light Emitting Diodes (LED) and Laser Diodes (LD), coming from the different stages of production and use of telecom equipment, including the qualification, the incoming inspection of the devices, the equipment testing phases and the field application. The major... of these results are related to GaAlAs devices, but some results on GaAs LED's and InP based LD's are also reported.

2. TECHNIQUES FOR FAILURE ANALYSIS AND TECHNOLOGICAL CHARACTERIZATION

Owing to the novelty of these devices, with respect to Silicon IC's, extended characterizations were performed for qualification purpose. Among the microscopical and microanalytical investigations performed the most effective was Scanning Electron Microscopy (SEM) examination of cross sectional specimens, as shown in figures 1 and 2. They were prepared by careful grinding and lapping single chips, removed from the package and tightly glued against a soft glass. The exposed surface, properly stained, clearly shows heterostructures and p-n homojunctions in GaAs. This information is necessary in failure analysis, which requires the combination of optical and infrared microscopy, conventional SEM and Electron Beam Induced Current (EBIC).

Figure 3 shows the images of a good GaAlAs laser, obtained by optical, infrared and scanning microscopy, and EBIC. Optical microscopy helps in detecting the most evident damage on chip, fiber or package. Infrared microscopy of the emitted light, or electroluminescence (EL)

[§]This work was partially supported by CNR under the "Progetto Finalizzato MADESS".

*Present address: Scuola Superiore di Studi Universitari e di Perfezionamento S.Anna-56100 Pisa - Italy

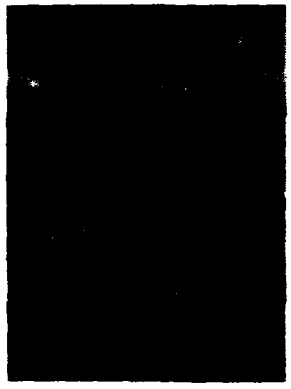


Fig. 1 SEM cross-sectional view of a GaAlAs oxide stripe laser.

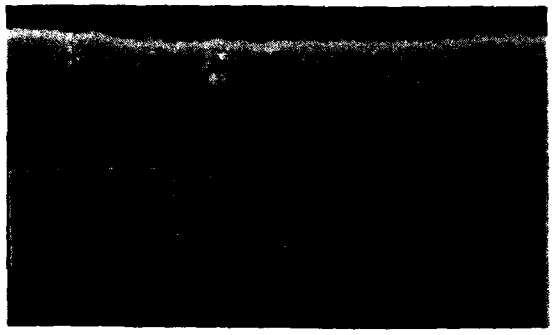


Fig. 2 SEM cross-sectional view of a InGaAsP buried crescent laser.

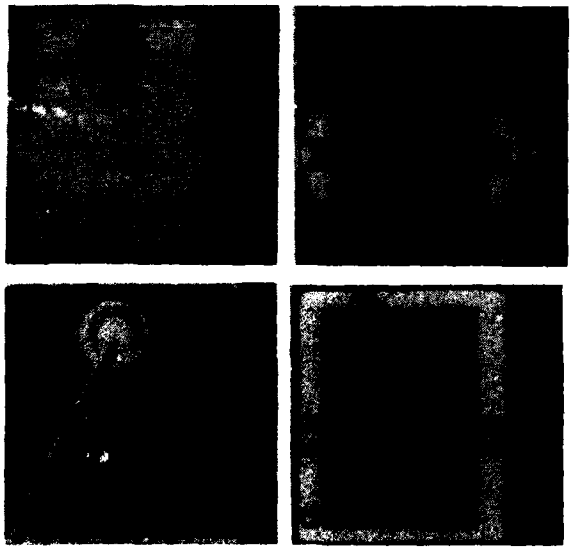


Fig. 3 a) Optical, b) EL, c) SEM and d) EBIC images of a good GaAlAs LD.

performed by means of a suitable converter applied to a conventional optical microscope, indicates possible emission anomalies. Unfortunately, it is shielded by the metallic contact layers and its resolving power is severely limited by lens aberrations at such a large wavelength. Conventional SEM gives high resolution images of device surfaces. EBIC technique is effective when the active region of the device lies near the top surface, because of the short penetration depth of the SEM electron beam. When applicable, it is a powerful tool in localizing high recombination sites in the whole area of the active layer, also out of the lasing region. The different distribution and thickness of the layers separating the active region from the top surface modulates the EBIC signal in a complicated way, so that a direct interpretation of EBIC images may be troublesome. However useful information is extracted by comparing the images of good and failed devices.

3. FAILURE ANALYSIS RESULTS

3.1. GaAlAs LD

The failure modes of many GaAs/GaAlAs oxide stripe LD's, coming from incoming inspection (I), qualification (Q), equipment production (P) and field application (F), were classified according to the optical power transmitted across the fiber, the photocurrent induced into the monitor diode and the I-V characteristics of the diode. Three groups of failure modes were identified, as shown in Table I.

TABLE I
Failure mode classification

	A	B	C
opt.pwr in fiber	none	none	reduced
monitor photocurr.	none	regular	reduced
laser IV charact.	degraded	regular	regular
percentage	18%	14%	68%
occurance	IPF	PF	QPF

Mode A was always due to evident destructive phenomena, such as the LD burn-out shown in

Fig. 4 (an electrical short circuit was measured), and melted or mechanically broken wires. The failure mechanism may be classified as an external overstress [5].



Fig. 4 SEM image of a burned GaAlAs LD.

Mode B was related to mechanical fiber shift (fig. 5) or misalignement, possibly caused by improper mounting.

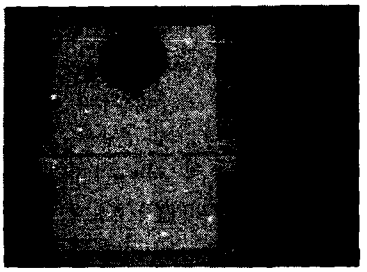


Fig. 5 Optical micrograph of fiber displacement

Mode C was the more frequent but also the more difficult to understand. It showed typical EL image with reduced emission, in agreement with electro-optical measurements (fig. 6), but a clearer understanding was given by EBIC investigation, which enabled us to identify three different situations: a) damage circumscribed to the mirror area (fig. 7a); b) formation of dark line defects (DLD) 6 into and around the lasing stripe (fig. 7b) and c) large dark areas outside the stripe region (fig. 7c). Mirror damage was usually present in cases b) and c), too.



Fig. 6 EL image of LD showing degraded emission at the facets.

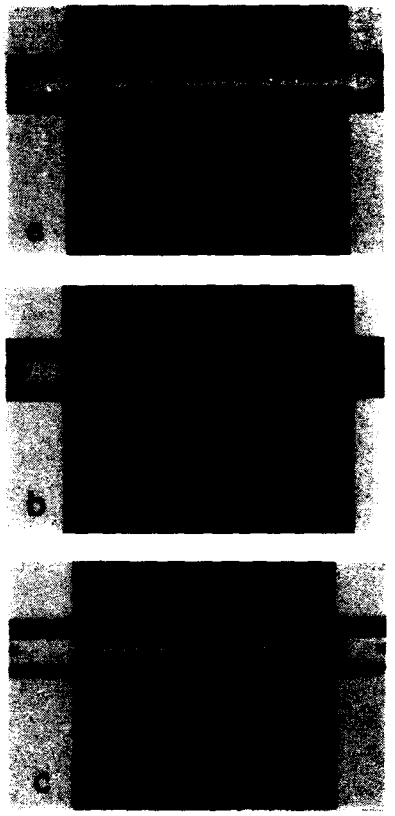


Fig. 7 EBIC image of degraded LD's. a) Only mirror damage, b) DLD in lasing stripe, c) large dark area of enhanced recombination out of the stripe.

Mirror damage is an effect of excess current injection. It was caused by erroneous operation in case a), and by the feedback circuit employed to keep the output power constant, in case b). No clear correlation between the images and the failure causes has yet been found for case c), although a mechanism similar to case b) could be assumed. In both cases the growth of the intrinsic defect reduces the radiative fraction of recombination, causing the monitor diode to measure a lower power emission and consequently to require more and

more current to restore the proper optical performance, until catastrophic damage of the highly stressed facets occurs.

3.2. GaAlAs LED

In spite of their simpler structure, LEDs are frequently more difficult to analyse: the lack of a double access to emission measurement (fiber and back-monitor diode in LDs, just fiber in LEDs) and the almost general impossibility of taking EBIC images of the devices having the active region at the bottom of the chip limited the available techniques to EL and optical microscopy.

Figure 8 shows the EL image of a LED whose failure mode was found in 75% of analysed devices and could be compared with mode C of Table I (optical degradation and correct electrical characteristics), and the associated failure mechanisms, that is the growth of large non-radiative regions, preferentially along defined crystallographic direction, seems to be related to metal-semiconductor interaction [7].

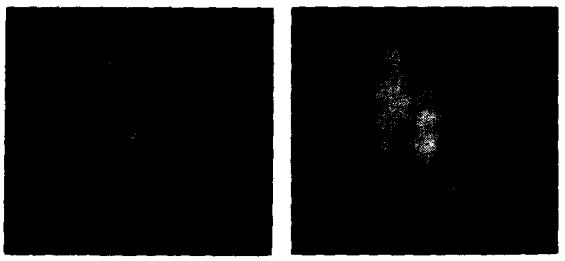


Fig. 8 Infrared images of a degraded GaAs LED. a) sum of emitted (EL) and reflected light; b) enlarged EL image of the emitting area, showing oriented defect structure.

3.3. InGaAsP LDs

For InP lasers a classification of failure modes like that of Table I cannot yet be drawn. Failure modes A and B, which are not related to the chip technology, have been identified (fig. 9), in 70% of the cases while the devices that show a situation like mode C are very difficult to analyse, due to the complicated structure involved. Work is in progress in order to extend previously described techniques to these optical emitters.

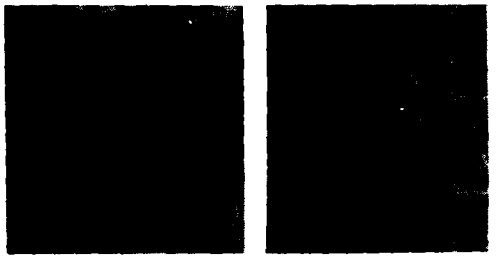


Fig. 9 a) SEM image of a burned InGaAsp LD. b) enlarged view of the burned area.

4. CONCLUSIONS

The techniques presented in this paper enable the failure mechanism affecting emitting devices for optical fiber communications to be identified.

Over 30% of the failures are due to the electrical overstresses and mechanical problems, in particular associated with the coupling between device and fiber. In the other devices, the damage of the facets of GaAlAs is a common failure mechanism, probably due to excess optical power in devices which failed in a very short time during testing, but couppled with DLD's in the active stripe for devices that lived for longer times in controlled conditions. Non radiative regions were also found in degraded LEDs.

REFERENCES

[1] Iwamoto, Y. and Fukinuki, H., IEEE J. Lightwave Techn. LT-3 (1985) 1005.

[2] Dean, B.A. and Dixon, M., Semiconductor and Semimetals 22 C (1985) 153.

[3] Douguet, A., Clavier, R. and Bleuzen, A. in Proc. 5th Int. Conf. on Reliability and Maintainability (Biarritz,1986) pp. 270-274.

[4] Fukuda, M., Kadota, T. and Uehara, S., Rev. ECL 34 (1986) 119.

[5] Sim, S.P., RObertson, M.J. and Plumb, R.G. J. Appl. Phys. 55 (1984) 3950.

[6] Petroff, P.M., Semiconductor and Semimetals 22 A (1985) 379.

[7] Zipfel, C.L., Semiconducotr and Semimetals 22 C (1985) 239.

TECHNOLOGICAL CRITICAL POINTS OF InGaAsP/InP 1.3 μm LASERS AS EVIDENCE OF 12000 HOUR CW OPERATING LIFE TESTS

Roberto DE FRANCESCHI, Michele LIBERATORE, Paolo MONTANGERO and Agnese PICCIRILLO

CSELT - Centro Studi e Laboratori Telecomunicazioni S.p.A. -
Via G.Reiss Romoli, 274 - 10148 Torino (ITALY)

Abstract. Electro-optical characteristics evolution of InGaAsP/InP lasers after 12000 hrs life tests with respect to their technological - crystallographic critical points is analyzed. In particular it was found the Indium die-attach is more reliale than is thought; in addition, through dynamic thermal resistance measurements, it is possible to screen out badly soldered devices. Gold diffusion was a degradation process always present even if it was not always the main one. Specific crystallographic defects result in influencing the ageing behaviour both directly favouring pulse threshold current increase and indirectly promoting thermal process like gold migration. Different SEM technics (SEI, EBIC, EDAX) and Auger analysis are employed to analyze degraded devices.

1 - INTRODUCTION

We discuss the results of 12000 hour life tests making correlations between electro-optical characteristics evolution and technological/crystallographic critical points.

The examined devices are 1.3 μm InGaAsP/InP lasers whose structure is shown in figure 1.

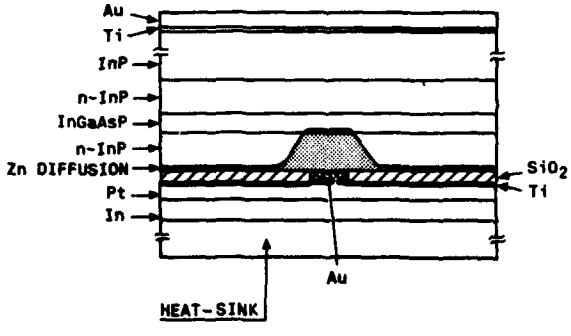


Fig. 1 - Schematic structure of tested InGaAsP/InP laser

They have been tested in CW operation at constant optical power (3 mW per facet) at 3 temperatures (25,35, 45°C).

Figure 2 shows the percentage variation of CW threshold current at 20°C for each device with time.

2 - DIE-ATTACH

All devices are mounted p - side down on a

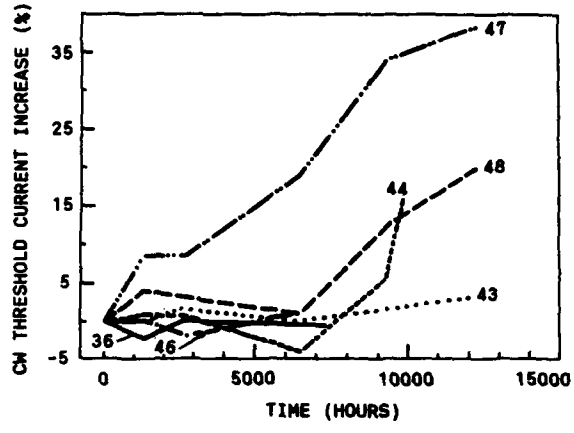


Fig. 2 - CW threshold current percentage variation at 20°C vs. ageing time

gold plated copper heat - sink with a thin film of indium solder. A careful observation showed the presence of some voids inside the solder, even if seldom located exactly below the stripe region. The presence of voids influence the device behaviour only when they are exactly located under the active region, that is when they are "hot voids" [1]. Dynamic thermal resistance measurements, based on voltage shift with junction temperature increase, makes it possible to separate the die-attach contribution and consequently to screen out badly soldered devices. In figure 3 the junction temperature increase vs. time of a badly

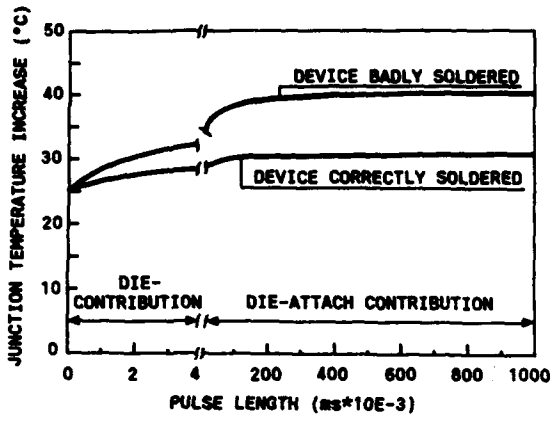


Fig. 3 - Comparison between dynamic thermal resistance measurements

soldered device is shown, compared to the typical behaviour of a properly mounted one.

Indium die-attach is considered not reliable mainly due to the formation of whiskers and intermetallic compounds with gold [2]. At the end of the life test all devices have been examined by SEM EDAX inspection and no evidence of these kinds of processes were found.

3 - CONTACT DEGRADATION

Referring to figure 2 it can be seen that the p - contact metallization consists of a gold alloy covered by titanium and platinum layers. Auger gold depth profiles of the contact stripe of two devices, one aged and the other a reference, showed gold migration inside the cladding layer in the former sample. Comparison between the two profiles is shown in figure 4. In the reference sample gold was found up to a depth of 0.25 μm (due to alloying process) while for a 12000 hour aged sample gold was found up to 0.75 μm, demonstrating that diffusion has occurred.

This degradation process was always present even if for some devices other specific mechanisms were stronger in determining ageing degradation.

4 - CRYSTALLOGRAPHIC DEFECTS

Devices 47 and 48 (both tested at 45°C) have shown a threshold current increase in CW operation and in pulsed operation. In addition device No.47 has shown a continuous increase of the threshold current starting from the beginning of the life test while the threshold current increase for device 48 started after about 6000 hours of operation life. SEM observations and EBIC analysis have shown two different crystallographic defects, responsible for the different ageing behaviour.

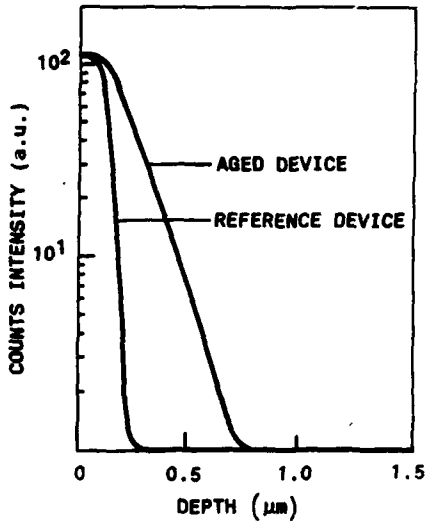


Fig. 4 - Comparison between Auger gold depth profiles in the stripe region

Figure 5 is an interference contrast image of the p - side of device No.47. The cross

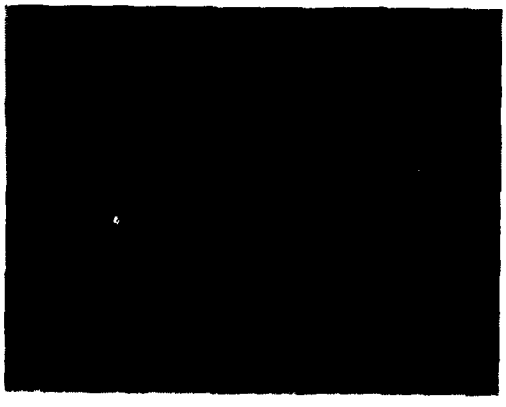


Fig. 5 - Cross - hatched pattern of misfit dislocations in device No.47

-hatched pattern of misfit dislocations is clearly visible. Figure 6 presents a pair of SEM images in SEI and EBIC modes respectively, showing the direct correspondance between the misfit lines in morphology and the dark lines in EBIC mode. These defective lines are in the [1 $\bar{1}$ 0] direction that is in agreement with previously published results on asymmetry of electro-optical properties of misfit dislocations in InGaAsP/InP heterostructures [3] [4].



a)



b)

Fig. 6 - SEM images showing the correlation between cross - hatched patterns in morphology and dark lines in EBIC mode.
a - Secondary electron image of the surface
b - Electron beam induced current image

As expected these misfit dislocations are located in the InP cladding layer as shown by a selective chemical etching of a part of this layer (figure 7). These misfit dislocations

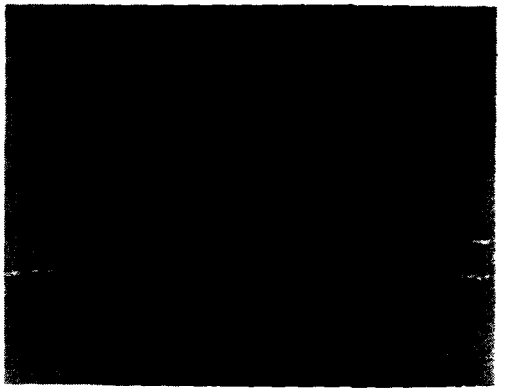


Fig. 7 - Device No.47 after selective chemical etching of a part of InP cladding layer, showing misfit dislocations are located in the cladding

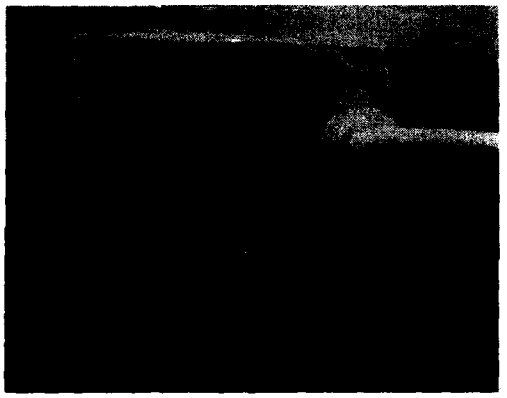
are caused by the known phenomena of the edge growth and this is consistent with the fact that device No.47 actually comes from the edge of the wafer as can be seen in figure 8. The low initial To value, 53°K compared with an average value of 65°K, is probably related to a higher over-barrier leakage, promoted by deep levels associated with the presence of



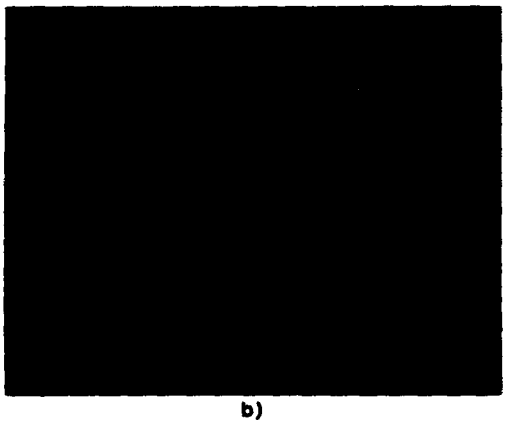
Fig. 8 - Front view of device No.47 coming from a wafer edge

misfit dislocations. The threshold current increase (CW and pulse) with ageing time can be due to mechanisms related to electronic processes and thermal processes, specifically propagation of misfit dislocations in the active layer with deep gold migration along particular paths, or contact degradation due to the uniform gold migration as discussed in the previous paragraph. The large increase in thermal resistance and in series resistance (ΔR_θ 30%, ΔR_S 20%) strongly support these hypothesis.

Figure 9 presents SEI and EBIC images of the p-side of device No.48 after detachment and metal removal. The EBIC image (35 KeV) shows a dark region, located across the stripe, whose size is about 10 x 15 μm. Figure 10 is an enlargement of this defective region. Comparison

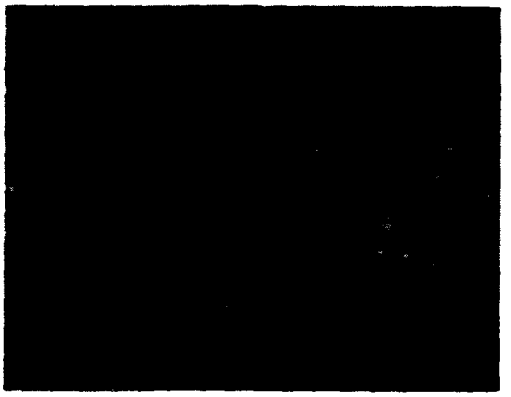


a)



b)

Fig. 9 - SEM images showing device No.48 after p-metal removal.
a - SEI image
b - EBIC image showing a dark area across the stripe



a)



b)

Fig.10 - Enlargements of the defective region. Both figures show the same area.
a - SEI image
b - EBIC image

between SEI and EBIC images clearly shows that there are no surface metal residues or other particles in the dark region. Consequently the dark area has to be related to bulk crystallographic defects and, considering size and geometry, it could be a dislocations-cluster which has propagated during epitaxial growth, starting from the InP substrate [5] [6]. The low initial To value, 43°K, supports this hypothesis. The large increase of the thermal resistance (ΔR_θ 64%) and CW threshold current

(ΔJ_{th} 20%) proves that the main degradation process has to have a thermal nature. As the die attach contribution in the thermal resistance increase was not significant, it was supposed that the contact region degradation was responsible for the observed behaviour. In fact SEM EDAX analysis of the dark region, after partial cladding removal with ion beam etching, brought to light gold segregated here, visible in figure 11. It is reasonable to think that the cluster of dislocations can become a preferred path for gold migration.



Fig.11 - SEI image of the same area of figure 10, showing gold segregated in the defective region, exposed by ion beam etching

5 - CONCLUSIONS

The results of 12000 hour life tests have shown that laser diodes behaviour is strictly connected to the technological choices. In particular the use of gold as first layer of p-side metallization is responsible for the main degradation for all devices; furthermore the choice of indium solder has proved to be reliable when used with TiPt metallization and in mild temperature conditions.

Even if InGaAsP/InP lasers are not basically considered susceptible to degradation due to crystallographic defects, we have found that misfit dislocations and clusters of dislocations, already present before the life test, can become the main cause in devices degradation.

ACKNOWLEDGEMENTS

The authors are grateful to F.Taiariol for helpful discussion.

This work has been partially supported by CNR progetto finalizzato "MADESS".

REFERENCES

[1] Yerman A.J., Burgess J.F., Carlson R.O. and Neugebauer C.A., IEEE Trans. On Comp.4 (1983) 473

[2] Mizuishi K.J., Appl.Phys.55 (1984) 289

[3] Yamazaki S., Kishi Y., Nakajima K., Yamaguchi A. and Akita K.J., Appl.Phys.53 (1982) 4761

[4] Yamaguchi A., Komiya S., Ueda O., Nakajima K., Umebu I. and Akita K., Gallium Arsenide and Related Compounds 1981, Conf.Ser. No.63 (1981) 161

[5] Stirland D.J., Hart D.G., Clark S., Regnault J.C. and Elliot C.R., J.of Crystal Growth 61 (1983) 645

[6] Franzosi P., Salviati G., Cocito M., Taiariol F. and Ghezzi C., J.of Crystal Growth 69 (1984) 388

TURN-ON DELAY TIME FLUCTUATIONS IN UNBIASED GAIN-SWITCHED AlGaAs/GaAs MULTIPLE QUANTUM WELL LASERS

E. H. Böttcher, K. Ketterer, and D. Bimberg

Institut für Festkörperphysik I , Technische Universität Berlin
Hardenbergstr. 36 , D-1000 Berlin 12 , Germany

The turn-on delay time jitter in unbiased gain-switched AlGaAs-GaAs multiple quantum well lasers is investigated in detail. The rms timing jitter is determined by measuring the transient fluctuations of the total emitted power using a fixed-time sampling technique with a picosecond temporal resolution. The turn-on jitter is found to decrease significantly with increasing pumping rate, particularly when the lasers are operated at an excitation level where only the first relaxation oscillation is emitted. The rms jitter of the emitted optical pulses decreases from about 20 ps just above the laser threshold to a value of 14 ps at the threshold for the appearance of the second relaxation oscillation. These results demonstrate that the accurate adjustment of the pumping rate is essential for a low jitter single pulse operation of the gain-switched lasers.

1. INTRODUCTION

Gain-switching of injection lasers by short electrical pulses is a convenient method for the generation of picosecond optical pulses /1/. There are numerous applications for such pulses in various areas like long range data transmission at gigabit per second rates /2,3/ and optical sampling /4,5,6/. Recently detailed experimental and theoretical investigations of the minimum pulse width which can be obtained for various types of lasers, cavity length etc. were reported /1/. The applicability of such a source of ultrashort optical pulses, however, depends not only on the pulse width and the pulse energy but to the same extent on the pulse-to-pulse timing jitter which has to be reduced to a sufficiently low level. In many experiments where repetitive pulses are utilized, like optical sampling, the timing jitter causes a degradation of the temporal resolution if its value is comparable or larger than the optical pulse width. For data transmission applications, it is evident that a low timing jitter is a requirement for a low error rate. The pulse-to-pulse timing jitter is an inherent phenomenon of a gain switched semiconductor laser. The spontaneous (random) character of the onset of the laser emission results in fluctuations of the turn-on delay time. Despite its large importance little work has been reported until now on experimental investigations of the size of the jitter, in particular for fully modulated lasers.

In this letter we focus our attention on the timing jitter of unbiased gain-switched lasers driven by short injection current pulses at comparatively low repetition rates. This case is important for all the applications for which the optical background signal between subsequent laser pulses has to be minimised like time resolved spectroscopy and the all optical boxcar system /6/. The lasers studied are AlGaAs-GaAs multiple quantum well devices which are particularly attractive for gain-switching applications because of their im-

proved modulation performance and lower threshold currents compared to conventional douple heterostructure lasers /7,8/. Experimentally, the statistical distribution of the turn-on delay time is measured by a modified version of a broad band sampling technique which was used for the study of partition noise in semiconductor lasers /9/. The observed probability density functions are fairly well described assuming a Gaussian distribution for the turn-on delay time. The root mean square (rms) value of the turn-on jitter is found to increase significantly if the pumping rate is decreased to the laser threshold. A rms jitter of up to 20 ps is registered in this case. With increasing pumping rate the turn-on jitter decreases and saturates with a corresponding rms value of about 8 ps. Our results demonstrate that, for low jitter optical single pulse generation, the pumping rate has to be adjusted to a value just below the threshold value for the appearance of the second relaxation oscillation in order to minimize the optical pulse width and the jitter simultaneously.

2. EXPERIMENTAL

In order to study the pulse-to-pulse timing jitter we measure the power distribution of the emitted laser radiation at a fixed sampling time. The repetitive optical transients are detected by an ultrafast GaAs-Schottky-photodiode (10% - 90% risetime $t_r \approx 10$ ps) and monitored by a sampling head ($t_r \approx 20$ ps). The sampling head is used as an ultrafast sample and hold circuit. It samples the instantaneous laser output with a picosecond time resolution and holds it for 33 us. The output of the sampling unit is processed by a multichannel analyser (MCA) operating in the sampled voltage analysis mode. An avalanche generator is used for the short electrical pulse excitation of the MQW lasers. The pulse amplitude is 40 V (at 50 Ω load) and the full width at half maximum (FWHM) of the pulse is 240 ps. The structure and the properties of the MQW lasers as well as their high frequency mounting have been described in detail elsewhere /8,10,11/. The lasers are mounted on a copper heat sink and the temperature is controlled by a Peltier cooler. All measurements are carried out at 296 K.

3. THEORY

In order to gain quantitative information on the timing jitter from the measured statistical data we have developed a model which is presented in the following. Assuming that the turn-on delay time t_d is normally (Gaussian) distributed its probability density function (PDF) is given by

$$P(t_d) = \frac{\exp\{-(t_d - \tau_d)^2/2\,\sigma_d^2\}}{\sigma_d \sqrt{2\pi}} \qquad (1)$$

τ_d and σ_d constitute the mean and standard deviation of t_d. Using the fixed time sampling method we measure the power fluctuations caused by the turn-on fluctuations. Therefore, the PDF of the power P(y) which corresponds to $P(t_d)$ should reflect the experimental results. An analytical expression for P(y) is derived with the reasonable approximation that the temporal waveform y(t) corresponding to the first RO has a Gaussian shape

$$y(t) = \exp\left(-(t - \tau_d)^2/2\,\sigma_p^2\right) \qquad (2)$$

y(t) is the normalized dimensionless time dependent power, $0 \le y \le 1$, representing a single laser pulse. In the

absence of any turn-on jitter, y(t) would be equivalent to the waveform displayed by the sampling oscilloscope. This waveform, however, is broadened due to the turn-on jitter resulting in a variance $\sigma^2 = \sigma_p^2 + \sigma_d^2$. can be determined directly from the temporal FWHM or the risetime of the waveform measured with the sampling scope.

P(y) is measured at a fixed sampling time $t = \tau_d - T$. Since $P(t_d)\, dt_d = P(y)\, dy$ it follows from (1) and (2) that P(y) is given by

$$P(y) = \frac{\exp\{-(2\sigma_d^2)^{-1}(\sqrt{-2\sigma_p^2 \ln y} - T)^2\}}{(\sigma_p)^{-1}\, 2\sigma_d\, y \sqrt{-\pi \ln y}} \qquad (3)$$

Experimentally, the samples are taken at the time $T_{1/2}$ within the leading edge of the first relaxation oscillation (RO) where the averaged emitted power reaches its half value of the maximum power. $T_{1/2}$ is given by

$$T_{1/2} = \sqrt{-2 \ln 0.5\, (\sigma_p^2 + \sigma_d^2)} \qquad (4)$$

It is evident from (3) and (4) that P(y) depends only on the ratio $k = \sigma_d / \sigma_p$.

In order to compare the experimental data with this model we have to take into account the additive source of amplitude noise due to the baseline fluctuation of the sampling oscilloscope which is the dominating vertical noise component compared with the amplitude noise of the pulse generators. The PDF of this vertical noise component y_v is independent and orthogonal to the timing jitter. It is measured by recording the output fluctuations of the sampling scope in the absence of a photocurrent signal. The PDF is found to be Gaussian with a standard deviation of 0.8 mV. The measured values y_m represent the sum of two independent random variables, $y_m = y + y_v$. Therefore, their PDF's are related by convolution

$$P(y_m) = P(y) * P(y_v) \qquad (5)$$

4. RESULTS

A typical experimental result for the power distribution is shown by the dotted curve in Fig.1. As shown by the solid curve in Fig.1 a good fit to the experimental data is obtained using the expression for $P(y_m)$ derived above. From the fits the value for $k = \sigma_d / \sigma_p$ is obtained. k is the only free fitting parameter. σ_d is determined using the additional result for the sum $\sigma_d^2 + \sigma_p^2$ given by the experimentally recorded waveform. The results for the total rms jitter σ_d determined in this way contain the inherent timing jitter of the gain-switched laser σ_{do} and that of the driving and measurement electronics σ_{de}. The latter component σ_{de} is analysed seperately applying the method outlined above. It is found that the core of this timing jitter is almost Gaussian distributed for the pulse generator in conjunction with the sampling scope, yielding an rms jitter σ_{de} of 3.5 ps. Assuming that the electronic jitter components are mainly caused by the sampling scope /12/, an upper limit for the inherent turn-on delay time jitter is obtained using the relation $\sigma_{do}^2 \simeq \sigma_d^2 - \sigma_{de}^2$.

The PDF data for minimal jitter at single pulse emission of the laser are shown in Fig.1. For this practiclly important case the intrinsic rms jitter is 14 ps. A further reduction of the amplitude of the voltage pulse by about 0.2 dB results in a rms jitter of about 22 ps. These results demonstrate how critical the adjustment of the pump pulse is,

if low jitter single pulse operation is searched for. Increasing the voltage amplitude about 2 dB above the laser threshold the timing jitter starts to saturate at a rms value of about 8 ps.

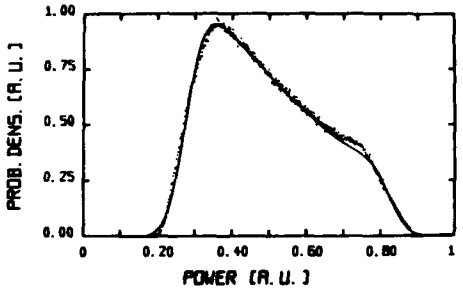


FIGURE 1

Probability density of the emitted light power measured at the fixed sampling time $T_{4/2}$. The dotted curve represent the experimental data for the MQW laser driven by short voltage pulses (FWHM = 240 ps). The excitation level is adjusted to the threshold of the appearance of the second RO. The solid curve is a fit to the data with k = 0.67 using the expression for $P(y_m)$ given in the text.

5. CONCLUSIONS

In summary, we have studied the turn-on delay time fluctuatins in unbiased gain-switched AlGaAs/GaAs MQW lasers at low repetition rates. A braod-band fixed-time sampling technique with picosecond time resolution was used to monitor the turn-on jitter. In order to evaluate the rms timing jitter from the observed statistical optical power distribution at a fixed sampling time, a simple model was developed which was found to describe the experimental data with high accuracy. In particular near the laser threshold, the timing jitter is extremely sensitive to the pumping rate. Therefore, the pumping rate has to be accurately adjusted to ensure low jitter single pulse operation of the laser.

REFERENCES

/1/ for recent reviews on picosecond pulse generation by gain-switching of injection lasers, see D.Bimberg, K.Ketterer, E.H.Böttcher, and E.Schöll, Int. J. Electronics 60, 23 (1986); P.-T.Ho, in Picosecond Optoelectronic Devices, ed. by C. H.Lee, (Academic Press, New York 1984) p.11

/2/ M.Danielsen, IEEE J. Quantum Electron. QE-12,657 (1976)

/3/ K.Y.Lau and A.Yariv, in Semiconductors and Semimetals, ed. by W.T.Tsang (Academic Press, Orlando 1985) Vol.22 Part B, p.71

/4/ T.Kanada and D.L.Franzen, Opt. Lett. 11,4 (1986)

/5/ D.L.Franzen, Y.Yamabashi, and T.Kanada, Electron. Lett. 23,289 (1987)

/6/ K.Ketterer, E.H.Böttcher, and D. Bimberg, Appl. Phys. Lett. 50,1471 (1987)

/7/ Y.Arakawa and A.Yariv, IEEE J. Quantum Electron. QE-21,1666 (1985)

/8/ K.Ketterer, D.Bimberg, M.Brezina, E.Schöll, G.Weimann, and W.Schlapp, Europhysics Conference Abstracts Vol.9H,126 (1985)

/9/ P.L.Liu and K.Ogawa, J. Lightwave Technology LT-2,44 (1984)

/10/ G.Weimann and W.Schlapp, Europhysics Conference Abstracts, Vol.8F, 363 (1984)

/11/ E.H.Böttcher, K.Ketterer, D.Bimberg, G.Weimann, and W.Schlapp, Appl. Phys. Lett. 50,1074 (1987)

/12/ W.L.Gans, Proc. IEEE 74,86 (1986)

CARRIER DENSITY AND INTER VALENCE BAND ABSORPTION IN InGaAs(P) LASERS

S.HAUSSER, E.ZIELINSKI, M.ASADA[+], H.SCHWEIZER, H.BURKHARD[*], E.KUPHAL[*]

4.Physikalisches Institut, Unversitaet Stuttgart
Pfaffenwaldring 57, D-7000 Stuttgart 80, FRG
[+] Fellow of the A. von Humboldt Stiftung,
on leave from Tokyo Institute of Technology
[*] Forschungsinstitut der Deutschen Bundespost
Am Kavalleriesand 3, D-6100 Darmstadt, FRG

A method is presented that allows the determination of the carrier density in semiconductor lasers above and below threshold. This method is based on a line shape analysis of the spontaneous emission. Corresponding gain values can be calculated from the carrier density. Temperature dependent measurements of the spontaneous emission and the differential quantum efficiency allow the determination of the optical losses at laser threshold. The result is a strong influence of inter valence band absorption on the threshold of 1.65 μm InGaAs lasers and an extremely weak influence on 1.3 μm InGaAsP lasers.

1. INTRODUCTION

Auger recombination and inter valence band absorption have been determined to be the main loss mechanisms, contributing to the low T_0-values in InGaAs(P)lasers [1]. However, the magnitude of the carrier density in the laser at threshold remains an open question, thus it is difficult to estimate the strength of the two loss mechanisms relative to each other. In the literature density values as high as $3 \cdot 10^{18}$ cm^{-3} [2] can be found. In contrast, from gain measurements on InGaAs epitaxial layers density values of $1.5 \cdot 10^{18}$ cm^{-3} can be deduced for a net gain of 100 cm^{-1} [3] which is necessary to overcome the mirror losses in a laser.

To overcome this discrepancy, we have performed measurements of the spontaneous emission spectra and differential quantum efficiency of various lasers operating at wavelengths between 1.65 μm and 1.3 μm. From the spontaneous emission spectra the carrier density can be determined by a line shape analysis, thus allowing the calculation of the temperature dependent material gain in the laser. This value equals the total optical losses as is confirmed by an independent measurement of the quantum efficiency.

2. MEASUREMENTS

The spontaneous emission of index guided lasers can be measured in sideward direction, perpendicular to the laser beam. In sideward direction, the spontaneously emitted light is only weakly influenced by stimulated recombination (whereas in forward direction the spontaneous emission spectrum is dominated by amplification and absorption effects). This holds for index guided lasers where the active layer is a small stripe, surrounded by higher band gap material.

The samples used were 1.65 μm InGaAs MS lasers [4] and a 1.3 μm InGaAsP BH laser. The light emitted in sideward direction was imaged onto a monochromator by a microscope objective and detected by a liquid nitrogen cooled Ge detector, using conventional lock-in technique. The measured spectra were corrected for the spectral characteristic of the experimental setup. The quantum efficiency measurements were done using a calibrated PbS detector. Care has been taken to image all light emitted by the laser in one direction onto the detector.

3. LINE SHAPE ANALYSIS

The spontaneous emission spectra were evalua-

ted by a line shape analysis using the following model:

- direct band to band recombination with k-selection
- constant interband matrix element
- Landsberg broadening [5].

Additionally, the amplification / absorption of the light in the 2μm wide active layer was taken into account using the expression for amplified spontaneous emission [6]:

$$I(E,w) \sim \frac{r_{spon}(E)}{g(E)} \left(e^{g(E)\cdot w} - 1\right)$$

where w is the width of the active layer (2 μm), g(E) is the optical gain / absorption and $r_{spon}(E)$ is the spontaneous emission rate.

These spectra were calculated at various temperatures and carrier densities. The carrier temperature was assumed to be equal to the bath temperature, as can be expected since the lasers were operated under pulsed conditions. This assumption is supported by the evaluation of the high energy tail of the spectra. The carrier density has been determined by a least squares fit of the theoretical spectra to the measured ones.

The result of such a fit is shown in fig.1 for the case of an InGaAs laser at 300 K and an injection current of 28 mA which is approximately equal to the threshold current. The agreement between experimental and theoretical spectrum is excellent at higher energies, supporting the assumption mentioned above. The deviations at low energies are caused by stray light of the laser beam in the cryostat and recombination via impurities. The carrier density in this case was determined to be about $1.25 \cdot 10^{18}$ cm^{-3} at T = 300 K.

Since the transition matrix element is known to be equal to 20.7 eV from absorption measurements [7] it is possible to calculate absolute gain spectra at a given carrier density, using the same theoretical model. In this case the maximum gain is about 190 cm^{-1}. To obtain the magnitude of the optical losses, the mirror losses of about 60 cm^{-1} for an active layer thickness of about 0.3 μm have to be subtracted from the maximum gain. Thus the total optical losses amount to about 130 cm^{-1}.

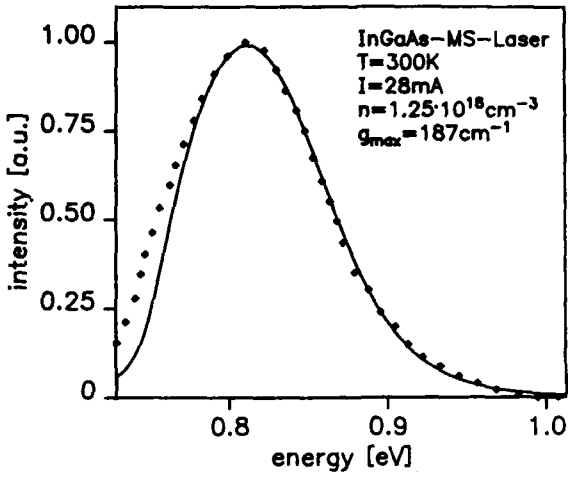


fig.1: experimental and calculated spontaneous emission spectrum of a InGaAs MS laser

4. RESULTS

4.1. Threshold Carrier Density

In fig.2 the result of the temperature dependent determination of the threshold carrier density is shown for the case of 1.65 μm and 1.3 μm lasers. The dots represent the values of the carrier density at laser threshold as determined from the measurements. The line represents the carrier density needed for zero material gain. The distance between the dots and the line thus indicates the additional carrier density needed to overcome the optical and mirror losses in the laser. This distance is nearly temperature independent in the case of 1.3 μm InGaAsP, whereas it is strongly temperature dependent for 1.65 μm InGaAs.

We also observe that the threshold carrier densities are generally lower in InGaAs than in InGaAsP. This results from the different band structure of the two materials, i.e. lower electron masses in InGaAs.

4.2. Optical Losses

The different behaviour of the two materials with respect to optical losses is shown in fig.3. The dots represent the laser threshold gain va-

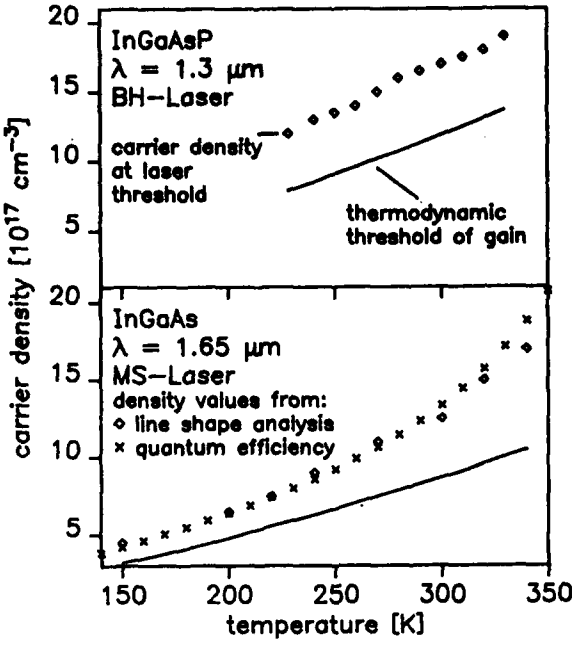


fig.2: temperature dependent carrier density values for InGaAs and InGaAsP. Dots are values determined from spontaneous emission measurements at laser threshold, crosses are calculated from the diff. quantum efficiency. The solid line is the threshold for zero gain.

lues calculated from the carrier densities. The crosses show the value of the total optical losses calculated from the quantum efficiency measurements by using the relation

$$\eta_D = \frac{1/L\ \ln(1/R)}{\alpha+1/L\ \ln(1/R)} = \frac{1/L\ \ln(1/R)}{\Gamma\alpha_{tot}}$$

where L is the length of the laser, R is the reflectivity of the mirrors, α are the optical losses, Γ is the optical confinement factor, α_{tot} are the total optical losses, including the mirror losses and η_D is the differential quantum efficiency. These total optical losses have to be equal to the threshold gain value. In fig.3 we observe good agreement between these two measurements in both cases. In 1.3 μm InGaAsP the optical losses turn out to be nearly temperature independent and in 1.65 μm InGaAs they show a a strong temperature dependence. This result corresponds to a nearly temperature independent quantum efficiency for 1.3 μm lasers. For the case of 1.65 μm lasers the quantum efficiency decreases at temperatures above 200 K.

The temperature independent part of the optical losses (indicated by the dashed lines in fig.3) is ascribed to mirror losses (60 - 100 cm^{-1}), free carrier absorption (20 - 40 cm^{-1}) [8] and scattering losses at the hetero interfaces. The temperature dependence of the free carrier absorption is negligible compared to the temperature dependence of the total losses. This temperature independent part of the losses has a magnitude of about 250 cm^{-1} for the 1.3 μm laser and 120 cm^{-1} for the 1.65 μm laser. This indicates a better quality of the interfaces in the case of the 1.65 μm laser.

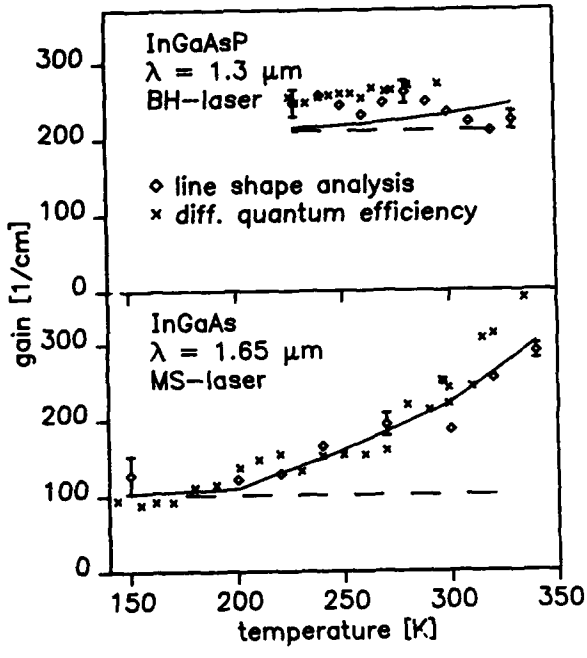


fig.3: temperature dependent gain values for InGaAs and InGaAsP lasers. The dots are determined from the line shape analysis, the crosses are calculated from the quantum efficiency measurements. The solid line is the calculated inter valence band absorption and the dashed line is the estimated temperature independent background absorption.

The temperature dependent part of the optical losses is ascribed to inter valence band absorption. That means that it is virtually nonexistent in 1.3 μm lasers. However in 1.65 μm lasers its magnitude is about 120 cm^{-1} at a temperature of 300 K (depending on the specific laser) and inter valence band absorption has a strong influence on

the threshold current. This is confirmed by theoretical calculations of inter valence band absorption.

5. CALCULATION OF INTER VALENCE BAND ABSORPTION

Again we used the model of a direct band to band transition between the heavy hole band and the split off valence band. Since this transition takes place at high k-values, we had to take into account the nonparabolicity of the heavy hole band. This was done by a linear interpolation of the band structures of GaAs, GaP, InP and InAs, given by Chelikowsky and Cohen [9]. In order to get good agreement between the calculations and the measurements, the usual inter band matrix element had to be divided by 16. The result of these calculations is also included in fig.3. The solid lines are the calculated inter valence band absorption values with the added temperature independent losses. The calculations can explain both, the material and the temperature dependence of the inter valence band aborption. The observed material dependence of inter valence band absorption is a result of the different band structures of the materials investigated. The strength of inter valence band absorption depends on the ratio of the split off gap Δ to the fundamental band gap E_g. Since this ratio increases with decreasing E_g in the InGaAs(P) material system, inter valence band absorption is stronger in the case of the 1.65 μm laser, even though it operates at a lower carrier density than the 1.3 μm laser.

6. CONCLUSION

In this paper we presented a method to determine the carrier density in semiconductor lasers. The results are threshold carrier densities of $1.2 \cdot 10^{18}$ cm^{-3} and $1.7 \cdot 10^{18}$ cm^{-3} at room temperature for InGaAs(P) lasers at 1.65 μm and 1.3 μm, respectively. These values are significantly lower than previous values published [2]. Temperature dependent measurements of the threshold carrier density and the differential quantum efficiency reveal a strong influence of inter valence band absorption in the case of 1.65 μm lasers, where its magnitude is about 120 cm^{-1} at room temperature. In the case of 1.3 μm lasers the magnitude of inter valence band absorption is negligible. The temperature and material dependence of inter valence band absorption was confirmed by calculations, taking into account the nonparabolicity of the heavy hole band.

ACKNOWLEDGEMENT

The financial support by the Deutsche Forschungsgemeinschaft under contract Pi 71/19 is gratefully appreciated.

REFERENCES

[1] Nelson, R.J., Dutta, N.K., Review of InGaAsP/InP Laser Structures and Comparison of Their Performance, in: Willardson, R.K., Beer, A.C. (eds.),Semiconductors and Semimetals (Academic Press, Orlando, 1985)

[2] Su, C.B. et al., Electron. Let. 18, 1108 (1982)

[3] Zielinski, E. et al., in: Haberland, D.H., Treusch, J., (eds.), Proceedings of the 5th General Conference of the Condensed Matter Division of the EPS, Europhysics Conference Abstracts, Vol. 9A

[4] Burkhard, H., Kuphal, E., IEEE J. Quantum Electron. QE-21, 650 (1985)

[5] Landsberg, P.T., Robbins, D.J., Sol. St. Electron. 28, 137 (1985)

[6] Cross, P.S., Oldham, W.G., IEEE J. Quantum Electron. QE-11, 190 (1975)

[7] Zielinski, E. et al., J. Appl. Phys. 59, 2196 (1986)

[8] Casey, H.C., Panish, M.B., Heterostructure Lasers, Part A (Academic Press, New York, 1980)

[9] Chelikowsky, J.R., Cohen, M.L., Phys. Rev. B 14, 556 (1976)

TWODIMENSIONAL MODEL FOR C^3- AND EXTERNAL CAVITY LASERS

J.P. Van de Capelle*, R. Baets, P.E. Lagasse
University of Ghent - IMEC
Laboratory of Electromagnetism and Acoustics
Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium

A twodimensional, self-consistent model for cleaved-coupled-cavity (C^3-) lasers is presented. The model is based on the Beam Propagation Method and includes the non-linearity of the medium : the longitudinal resonances are calculated in consistency with the optical power, which affects the refractive index. The method reveals some phenomena which cannot be found by means of a onedimensional model.

1. INTRODUCTION

During the last years people have searched for reliable, tunable and fast switchable lasers. It has been noticed that external optical feedback may increase the mode selectivity of a laser. A number of models have therefore been developed for better understanding of these lasers. Many models are based on (spatially independent) rate equations, and only few models include one spatial dimension.
To our knowledge this is the first report of a model for compound cavities including two dimensions, namely the longitudinal dimension and one of the transverse dimensions. Although this model is relatively complex, it has the advantage of taking into account a large number of physically relevant interactions between the optical field and the refractive index distribution. Because of this accurate description the model can reveal some phenomena which cannot be found with a more simple model.

2. DESCRIPTION OF THE MODEL

The model includes one lateral mode, which sasatisfies a twodimensional, scalar, non-linear Helmholtz equation. This equation is solved by means of the Beam Propagation Method, which is used to propagate the field backward and forward through the whole (compound) cavity.

The optical field intensity alters the complex effective refractive index, through the stimulated emission and the plasma effect. This is similar as for single cavity lasers [1].
At the waveguide discontinuities the propagating fields are partly reflected and partly transmitted according to the Fresnel laws. In this way the method determines the resonant solutions of the compound cavity. In order to obtain convergence, a special iteration procedure was developed. Starting from an initial guess of the field distribution, the power and the wavelength, these quantities are adapted until a fully self-consistent solution is obtained.. First we iterate on the power and the mode profile. This is done by propagating the field backward and forward, keeping the wavelength fixed. After each roundtrip the phase of the beam is normalised. When a number of roundtrips have been done the mode profile and power converge to a stable solution. Then we need to adapt the wavelength so as to satisfy the phase resonance condition. As this condition is coupled to the field intensity, through the non-linearity of the laser medium, we have to restart the intensity iteration. This procedure is repeated until a fully self-consistent solution is found, satisfying all boundary conditions. A similar iteration procedure is used in one-dimensional models [2], [3].

* Supported by the Belgian National Fund for Scientific Research (NFWO).

3. EXAMPLES

The model has been applied to an index guided C^3-laser, for which the gap width was chosen equal to an integer multiple of half a wavelength, and an empirical transverse diffraction and coupling loss was included. The lateral diffraction loss in the gap is accurately included in this method by means of the BPM-method. Two cases will be studied in detail. In a first case we fix the laser current at I_1 = 65mA and the modulator current is varied. In a second case the modulator current is fixed at 30mA and laser current is varied. For the first case, the fig. 1-2 show the output powers for each mode from the laser mirror (P_0) and the modulator mirror (P_L) respectively. The inset in fig. 1 depicts the situation.

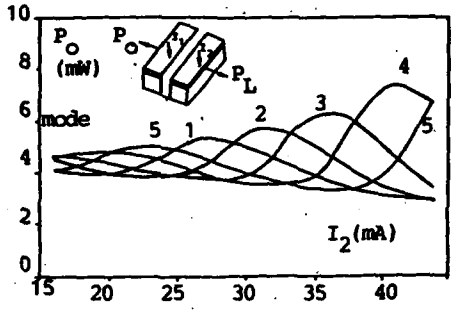


FIGURE 1

Output power P_o(mW) from the laser section.

We will restrict ourselves to a situation for which the first section definitely remains the dominant section. As a consequence, we only need to consider five consecutive modes of this section, since other modes will be perfect repeat modes of these five.
If the modulator current is increased, the mode powers behave oscillatory, the dominant mode having the largest output power P_o. The fig. 1 therefore indicates the wavelength tuning which occurs as the modulator current is altered. This wavelength tuning is a consequence of the anti-guiding effect which superimposes itself on the built-in refractive index. If the modulator current is increased, the refractive index of that section slightly decreases. This results in a shift of the reflection coefficient, seen by the first section, towards shorter wavelengths, and hence the threshold current and quantum-efficiencies of the different modes of the laser section are altered. From fig. 1 it is seen that the maximum outputpower P_o of each mode is increasing with increasing modulator current. This power increase is strongly influenced by the mutual coupling parameter
$a_c = S_{12}S_{21}/(S_{11}S_{22})$, where S_{ik} are the S-parameters of the gap [3]. The larger $|a_c|$ the stronger the coupling between the two sections and the larger this power increase in P_o will be. From fig. 1 it can also be seen that a maximum output power of one mode (nearly) coincides with the crossing of the power-current curves of two other modes. This indicates that the dominant mode then coincides with a reflection maximum, while two other modes are situated symmetrically around the dominant mode, 'observing' the same reflection from the gap and the modulator. This situation therefore corresponds to a situation of optimal mode rejection.
Fig. 2 depicts the ouput power P_L from the modulator section. The inset in fig. 2 shows the lateral far field distribution (normalised intensity and phase) of the dominant mode, for I_2 = 44mA. Comparing fig. 1 and 2 one observes that the maxima of P_o and P_L do not perfectly coincide : an offset of nearly 3mA between these maxima may occur. This is important, because the output power P_L from the modulator is usually detected to monitor the modulator current to a point of optimal spurious mode rejection. From fig. 1 and fig. 2 it is however easily seen that the crossings of the different curves do nearly occur for the same modulator current.
From fig. 1 and fig. 2 we can furthermore see that the mode rejection, at points of optimal biasing, increases with increasing modulator current. This is a well-known phenomenon, already reported in [4].
Fig. 3 shows the wavelength (λ) of the diffe-

rent modes. By means of an analytic approximation it can be shown that the wavelength variation with modulator current is proportional to [3]

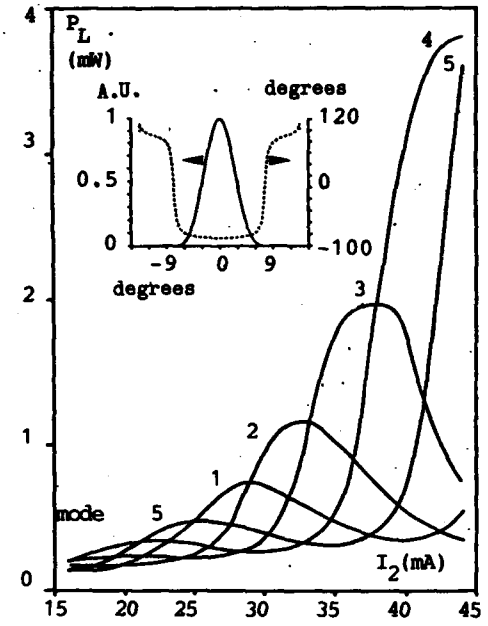


FIGURE 2
Output power P_L(mW) from the modulator section for the different longitudinal modes (1-5) (I_1 = 65mA).

$$\frac{1}{\lambda}\frac{d\lambda}{dI_2} \sim -(1 + b_{eff}^2)\,\frac{1}{|u|\rho}\,\frac{\partial|u|}{\partial\theta} \qquad (1)$$

in this formula ρ is the modulator roundtrip gain, b_{eff} is an effective anti-guiding parameter, $|u|$ is the normalised reflection seen by the laser section from the modulator and the gap and θ is the modulator roundtrip phase. From (1) it is seen that the wavelength variation with modulator current becomes zero as the mode becomes aligned with a reflection maximum. Comparing fig. 1 and fig. 3 it can indeed be seen that the maxima of P_o coincide with points for which the anologue wavelength tuning becomes zero. According to (1) the maxima of P_o thus coincide with a reflection maximum.
In a second example we have fixed the modulator current I_2 at 30mA and we changed the laser current. The resulting output power P_o (full line) and $P_o + P_L$ (dashed line) has been drawn in fig. 4. The main effect of the increase

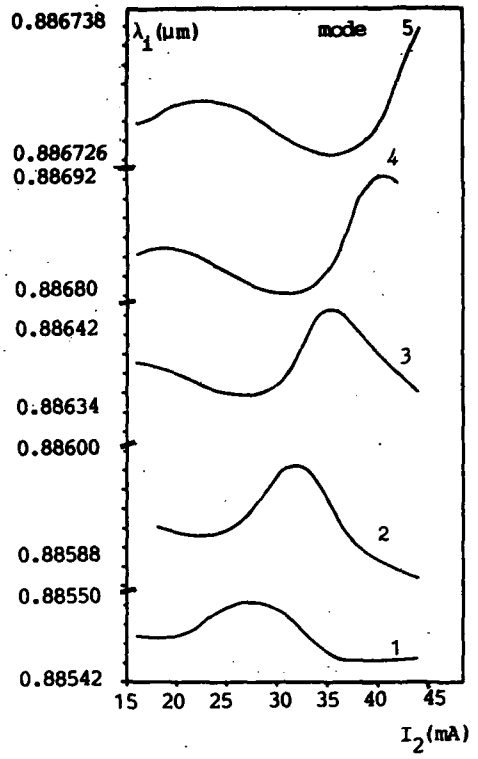


FIGURE 3
Wavelengths of the different longitudinal modes (1-5) (I_1 = 65mA). The wavelength becomes stationary as the mode becomes aligned with a reflection maximum or a reflection minimum.

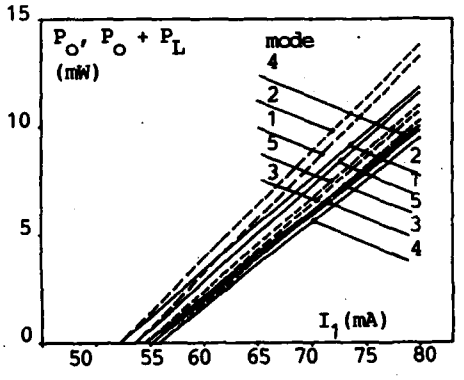


FIGURE 4
Output power P_o from the laser section and total output power $P_o + P_L$.

in laser current is a linear increase in output power. This indicates that the electron density within each laser section is nearly clamped. The influence of the power on the optical feedback from the gap and the second section upon the first section is nearly negligible. This is simply a consequence of the fact that the modulator section is biased near transparancy (see also [4]).

4. CONCLUSION

We have presented a twodimensional for C^3-lasers and external cavity lasers. Some relevant phenomena with respect to the stabilisation of the output of the laser have been revealed by this model. We found that one should be careful in detecting the modulator output for monitoring the laser current to a point of optimal mode rejection.

ACKNOWLEDGEMENT

The authors would like to thank P. Vankwikelberge for useful discussions on this topic.

REFERENCES

[1] R. Baets, J.P. Van de Capelle, P.E. Lagasse, 'Longitudinal Analysis of Semiconductor Lasers with Low Reflectivity Facets', IEEE J. Quantum Electron., Vol. QE-21, No. 6, pp. 693-699, 1985.

[2] J.P. Van de Capelle, R. Baets, P.E. Lagasse, 'Multi-longitudinal-Mode Model for Cleaved-Coupled-Cavity Lasers', IEE Proceedings Part J, Vol. 134, No. 1, pp. 55-64, 1987.

[3] J.P. Van de Capelle, R. Baets, P.E. Lagasse, 'Multi-Longitudinal Mode Model for Cleaved Coupled Cavity Lasers, Part II : Calculation Results and Discussion', to be published in IEE Proceedings Part J.

[4] C.H. Henry, R.F. Kazarinow, 'Stabilisation of Single Frequency Operation of Coupled Cavity Lasers', IEEE J. Quantum Electron., Vol. QE-20, No. 7, 733-744, 1984.

P4.1.14

SIMPLE MODEL FOR SEMICONDUCTOR LASER END FACET REFLECTIVITY

Gian Paolo BAVA, Andrea BIANCO

Dipartimento di Elettronica, Politecnico di Torino
C.so Duca degli Abruzzi, 24, 10129 Torino, Italy

Ivo MONTROSSET

Dipartimento di Ingegneria Biofisica ed Elettronica, Università di Genova
Via All'Opera Pia 11a, 16145 Genova, Italy

The mode reflectivity of narrow stripe double heterostructure semiconductor laser is evaluated using a simplified approach. It is based on a mode matching technique in the Fourier Transform domain for a suitable equivalent structure. The method allows to compute the reflectivity by the evaluation of an integral in the spectral domain and to obtain a closed form analytical expression for the radiation pattern.

1. INTRODUCTION

The end facet reflectivity of semiconductor laser devices has been extensively treated in the literature mainly for a planar structure by several authors |1,2,3,4,5, etc| with different techniques. The three dimensional waveguide case, whose interest is becoming of increasing importance owing to the use of very narrow strip devices, has been so far considered only by two authors |6,7|.

The relevance of the present problem is considerable not only for the evaluation of mirror effects on laser oscillators, but also as regards the computation of the radiation pattern |2| (a well known characterization parameter) and for the design of laser amplifiers antireflection coatings |5|.

The previously reported solution |7| of the problem under analysis, being based on a rigorous formulation, is rather cumbersome. In this paper a simplified technique is presented; it requires a small amount of computer time and, as a consequence, can be easily used for device otpimization. Extensive computation for the quasi-TE modes end facet reflectivity have been carried out; the evaluation requires only the computation of an integral in the spectral domain. The formulation allows also to obtain in closed analytical form an expression for the far-field radiation.

2. FORMULATION OF THE PROBLEM

The model used for our formulation of the problem is shown in Fig. 1 (in the actual structure $\delta = 0$). The interface between the waveguide and the air cannot be in the general described as a boundary between two uniform media owing to the difference of mode characteristics in the two regions.

The difficulties of the mode matching formulation, for the evaluation of the reflection coefficient for an incident mode, are connected with this point and with the presence of the continuous spectrum in both structures.

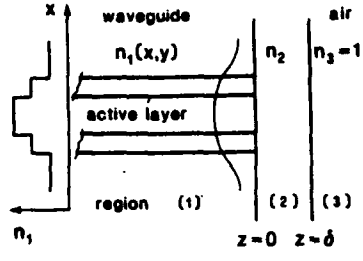


Figure 1

Two-dimensional model of the structure used for the problem formulation: (1) waveguide region, (2) "ficticious" transition region, (3) air.

The reason for the large number of formulations of the two-dimensional problem is then mainly related with the degree of approximations introduced and with the possibilities to handle the continuous spectrum:

- neglecting the continuous spectrum in the waveguide region |1|
- discretization of the continuous spectrum in the air and in the waveguide region |3,4|
- rigorous formulation using the mode sets of the two regions and construction of iterative solutions |2,5|

The computational difficulties are greatly increased for the three-dimensional problems for:

- the approximate evaluation of the waveguide modes (effective refractive index approximation)
- the vectorial formulation; the two-dimensional problem is strictly scalar
- the bidimensionality of the set for the expansion of the continuous spectrum.

The first approach to this case |6| was obtained neglecting the continuous spectrum in the waveguide region. The most recent one |5|, based on a rigorous extension of the work in |2|, is relatively cumbersome from a numerical point of view.

Our formulation is based on the quasi-TE and quasi-TM approach for the mode fields in the waveguide and in the air, on a proper matching in the Fourier Transfor (F.T.) domain between the incoming reflected and transmitted field and it takes into account mode conversion and the radiation spectrum in the waveguide region.

In order to be allowed to introduce a Fourier expansion for the modes we have followed the idea |8| to introduce a fictitious homogeneous layer of refractive index n_2 and "zero thickness" that minimize the reflection and the field deformation for the incoming field at the waveguide interface.

In the Fourier domain (ξ,η) the total electric field for quasi-TE modes is written in the form

$$E_1(\xi,\eta)=[1+R(\xi,\eta)]\ E_\mu(\xi,\eta)\cong E_2(\xi,\eta)=E_3(\xi,\eta) \quad (1)$$

where E_n stands for the y component of the electric field in region n, E_μ and R for the F.T. of the incoming mode field and for the proper reflection coefficient of its (ξ,η) components.

The continuity is set on the transverse magnetic field $H_x=H$ whose incoming, reflected and transmitted components can be computed using the relation

$$j\omega\mu_o \frac{\partial^2}{\partial z} H_x \cong \left(\frac{\partial^2}{\partial x^2} + K_o^2 n^2\right) E_y \quad (2)$$

In region (2) and (3) this relation becomes

$$H(\xi,\eta) = Y_o \frac{(n_j^2K_o^2 - \xi^2)\ (1-R)}{(n_j^2K_o^2-\xi^2-\eta^2)^{1/2}} E_\mu(\xi,\eta) \quad (3)$$

and in the waveguide is written in the form

$$j\omega\mu_o \frac{\partial}{\partial x} H_x - \left(\frac{\partial^2}{\partial x^2} + K_o^2 n_2^2\right) E_y = K_o^2(n_1^2(x,y) - n_2^2)\ E_y \ . \quad (4)$$

In our approximate formulation we have choosen n_2 in such a way to set to zero the average value of the second member, in (4).

This choice gives zero reflection in the limit of a first order approximation as can be shown both using coupled mode theory |9| or the results of |2| in the two dimensional case.

With this approximation both the Fresnel reflection coefficient for incident quasi-TE field on a plane interface between two media with refractive index n_2 and n_3 or the matching in the spectral domain between the Fourier components of the field at the interface give

$$R(\xi,\eta) = \frac{(n_2^2K_o^2-\xi^2)\ \zeta_3^* - (n_3^2K_o^2-\xi^2)\ \zeta_2}{(n_2^2K_o^2-\xi^2)\ \zeta_3 + (n_3^2K_o^2-\xi^2)\ \zeta_2} \quad (5)$$

where $\zeta_j^2=(n_j^2 K_o^2 - \xi^2 - \eta^2)$.

The evaluation of the reflection coefficient for the incident mode can be computed immediately by using the mode ortogonality in the waveguide

$$R_\mu = \iint (n_2^2 K_o^2 - \xi^2)\ R(\xi,\eta)\ |E_\mu|^2 d\xi\ d\eta \quad (6)$$

where the following normalization condition holds

$$\iint (n_2^2 K_o^2 - \xi^2)\ |E_\mu|^2\ d\xi\ d\eta = 1 \quad (7)$$

3. COMPARISON WITH OTHER FORMULATIONS

For what concern the three-dimensional case the formulation in |6| can be obtained by imposing $R(\xi,\eta) = R_\mu$ and neglecting $\partial^2/\partial x^2$ in (2) that limits the validity of that formulation to wide strips.

It is very difficult to compare analytically our formulation with the most rigorous in |7|. For this reason a comparison has been carried out with |2|. The leading term of the reflection coefficient gives exactly the same analytical expression of our approach.

The only consideration about |7| is that the effort to give a rigorous vectorial characteristic for the field in the air coicide with the quasi-TE approximation used in our formulation.

4. NUMERICAL RESULTS

An extensive comparison with the results found in the literature has been carried out. Due to the relatively small number of results for the three-dimensional case |10| a particular attention has been given to test our formulation for two-dimensional case.

In Fig.2 the results of Ikegami |1| for the amplitude of the reflection coefficient of the TE_o mode are compared with ours.

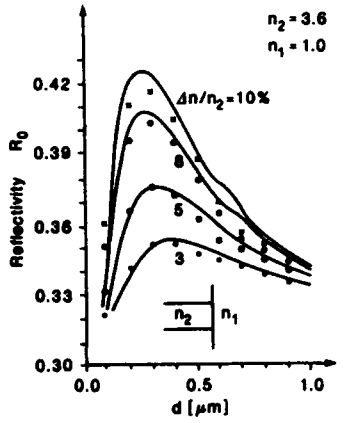


Figure 2

TE_o mode power reflection coefficient for a planar symmetric DH laser with active layer thickness d and n=3.60. Comparison between Ikegami |1| results (——) and ours for some values of the refractive index steps at $\lambda = 0.85$ μm.

In order to test also the phase we compared the Rozzi results in |3| for a single mode slab waveguide 0.6 um thick with n=3.61 in the core and 3.40 in the cladding. For $\lambda = 0.9$ μm we obtained $|R_o|^2 = 0.6216$ and $\phi = 2.15°$ instead of 0.622 and 2.92°.

More estensive results were found in |4| and are compared in Fig. 3(a-b) with ours.

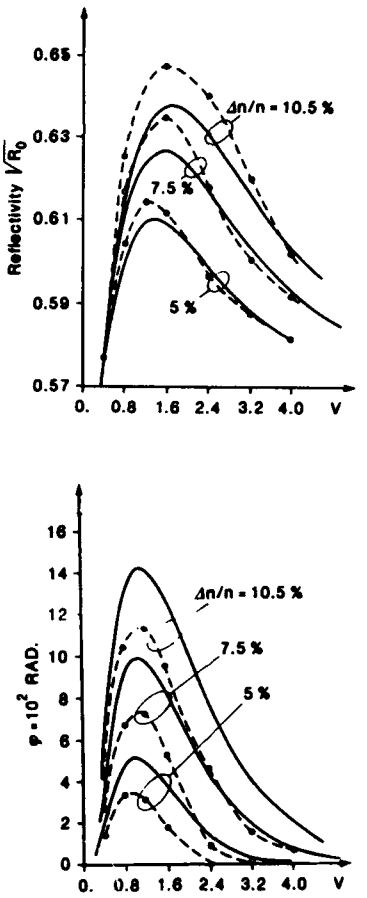


Figure 3

Reflection coefficient for the same structure in Fig.2 as a function of $V = K_o n \Delta n d$. Comparison between the results in |4| (——) and ours: (a) magnitude (b) phase.

For the three dimensional case the comparison with the numerical results in |10| is shown in Fig.4 and no comparison with the results of the simpler formulation in |6| has been reported due to their limitation to wide strips.

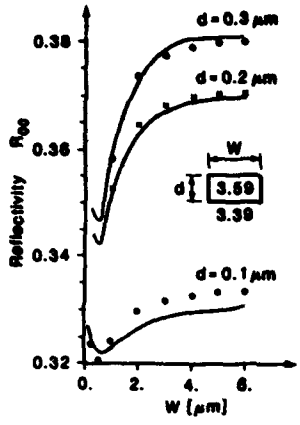


Figure 4

Quasi-TE_{00} mode power reflection coefficient of a BH laser. Comparison between the results in |10| (——) and ours.

Previous comparisons show a maximum error around 2.5% and 5% for the power reflectivity respect to the results in |2| and |4|.

The discrepancy on the phase is greater and it reaches a maximum of 0.03 rad.

The results for the three-dimensional case are quite good and show for narrow strips a decrease of the reflectivity connected with the increased TM character of the incoming quasi-TE field.

5. CONCLUSIONS

In this communication a simplified formulation for the evaluation of the end facet reflectivity has been presented. Its validity has been checked with the most relevant results in the literature for the TE mode of two- and three-dimensional planar waveguide. The formulation of the problem in the spectral domain allows a straighforward application to the evaluation of the radiation pattern and to the design of antireflection coating for laser amplifier.

ACKNOWLEDGMENTS

The authors are grateful to V.Ghergia and G.Destefanis for stimulating discussions and to CSELT for the partial support of this work.

REFERENCES

| 1| Ikegami T., IEEE J. of Quantum Electron., QE-8 (1972) 470.
| 2| Lewin L, IEEE Trans. Microwave Theory Tech., MTT-23 (1975) 576.
| 3| Rozzi, T.E., In't Veld, G.H., IEEE Trans. Microwave Theory Tech., MTT-28 (1980) 61.
| 4| Pudensi, M.A.A., Ferreira, L.G., J.Opt. Soc. of Am., 72 (1982) 126.
| 5| Kaplan, D.R., Deimel, P.P., AT&T Tech. J., 63 (1984) 857.
| 6| Kardontchick, J.R., IEEE J. Quantum Electron., QE-18 (1982) 1279.
| 7| Hardy, A., J. Opt. Soc. of Am., part A, 1 (1984) 550.
| 8| Vassallo, C., Electronics Lett., 21 (1985) 333.
| 9| Marcuse, D., Theory of Dielectric Waveguides, (Academic Press, New York, 1974.
|10| Handelman, D., Hardy, A., Katzir, A., IEEE J. Quantum Electron., QE-22 (1986) 498.

P4.1.15

A TUNABLE TWIN-STRIPE DISTRIBUTED-FEEDBACK LASER MODEL

M Federighi and T Kumar*

Marconi Italiana, 1 via A Negrone, 16153 Genova Cornigliano, Italy

* GEC Research Limited, Hirst Research Centre, East Lane, Wembley, UK

1. INTRODUCTION

There is a growing interest in tunable distributed-feedback (DFB) lasers, because of their prospective importance in various applications. Tunability is currently achieved by dividing the laser cavity into an active region and a tuning region [1,2]; a change in current through the tuning region changes the carrier density in the active region, thereby changing the effective refractive index and the Bragg wavelength. A double-channel planar buried heterostructure (DC-PBH) has been used in both references [1] and [2].

Most theoretical treatments of DFB lasers have concentrated on broad-area models [3-5], where modifications in the longitudinal structure such as a $\pi/2$ phase slip in the grating have been used in order to achieve single longitudinal-mode operation. The lateral structure of the devices has only been considered in the study of DFB laser phase arrays [6], using the coupled-modes theory which is not valid for strongly coupled stripes. However, the lateral field and carrier density profiles of the device influence the longitudinal spectrum through an alteration of the effective refractive index and of the Bragg wavelength [7]. This effect must be taken properly into account, particularly in the design of tunable DFB lasers, since alterations of the lateral geometry of the device can significantly affect its frequency spectrum.

We present in this paper a more general subthreshold model of a twin-stripe DFB laser, taking into account the current spreading in the passive layer and the lateral diffusion of carriers in the active layer. The purpose of this research is to investigate the effect of different (symmetrical and asymmetrical) pumping conditions on the longitudinal spectrum and therefore on the laser tunability. The electrical part of the model has been treated by solving a two-dimensional Laplace equation consistently with the carrier diffusion equation in the lateral direction, whereas the corresponding field has been calculated using the beam-propagation method (BPM).

A short description of our model is presented in Section 2, and the results in Section 3; Section 4 contains our conclusions.

2 THE MODEL

a) Electrical part

The model used in this analysis is substantially the same described in Reference [8]. We have assumed that the highly doped p-type "cap" layer does not contribute significantly to the current spreading, and that the n-type passive layer and the substrate can be replaced by an equipotential contact at the active layer interface; also, that the two stripes are ohmic and equipotential contacts. No variations in the longitudinal direction are taken into account.

For given applied potentials on the two stripes, the current distribution in the p-type passive layer is determined by solving the two-dimensional Laplace equation $\nabla^2 V=0$, subject to the insulated boundary conditions[6] $\nabla V.n=0$.

The current density injected into the active layer is given by:

$$J(x)\Big|_{y=d} = -\sigma \nabla V(x)\Big|_{y=d} \qquad (1)$$

where σ is the conductivity in the passive layer. J(x) acts as the source of the injected carriers, n(x), in the active layer; the carrier density distribution n(x) can be obtained from the diffusion equation:

$$D_{eff}\frac{d^2n(x)}{dx^2} - Bn^2(x) = \frac{-J(x)|y=d}{et} \qquad (2)$$

where D_{eff} is the effective diffusion constant [9], Bn^2 is the bimolecular recombination term and t is the active layer thickness.

Using Joyce's expressions for the Fermi integrals [10], the potential across the p-n junction can be written as a function of the carrier density n(x); the above equations can thus be solved self-consistently.

b) Optical part

In our analysis of the optical field inside the laser cavity, we have used basically the model of Reference [7] with some modifications. We have assumed the solution of the scalar wave equation to be a superposition of a forward travelling wave ϕ^+e^{ikz} and a backward travelling wave ϕ^-e^{ikz}, where $\phi^{\pm}(x,z)$ are slowly varying functions of z and the wave vector k is related to the grating period via the Bragg resonance condition. Following Reference [11] we have expressed the grating-induced perturbation in the dielectric constant ε as h(K/k)cos(2kz), K being the coupling coefficient of the grating.

To take the lateral carrier profile n(x) into account, we have followed Reference [12] and assumed a linear dependence of gain on carrier concentration, g(x)=an(x)+b, and a linear "antiguiding" perturbation -Ran(x). Including a built-in lateral index step $\Delta N(x)$ as well, the total perturbation $\Delta\varepsilon$ in the dielectric constant can be written as [12]

$$\Delta\varepsilon = 2\Gamma N_a\Delta N(x) - (\Gamma/k_o)Ran(x)N_a - i\Gamma N_a g(x)/k_o + i(1-\Gamma)N_p\alpha_c/k_o \qquad (3)$$

where N_a and N_p are the refractive indexes of the active and passive layers, Γ is the transversal confinement factor, k_o is the free-free-space wave vector and α_c is the absorption coefficient in the passive layer.

Substituting $\phi^+e^{ikz}+\phi^-e^{-ikz}$ into the scalar wave equation, dropping the negligible terms $\partial^2/\partial Z^2$ and simplifying, we obtain the coupled equations

$$\frac{\pm\partial}{\partial Z}\phi^{\pm} = i\left[\frac{\Gamma\omega}{v_g} + \frac{1}{2k_oN_{eff}}\frac{\partial^2}{\partial x^2} + \frac{k_o}{2}\Delta\varepsilon\right]\phi^{\pm} + iK\phi^{\mp} \qquad (4)$$

that differ from the similar equations used in broad-area models because of the presence of the "lateral" (x-dependent) operator added to $\Delta\omega/v_g$. $\Delta\omega$ is the difference between the light frequency ω and the Bragg frequency ω_B, and v_g is the group velocity; N_{eff} is the effective refractive index. We have solved Equation (4) by use of the beam-propagation method (BPM); details may be found in Refereces 7 and 12.

Above threshold, a term $-g(x)|\phi|^2$ representing the stimulated recombination must be added to the right-hand side of the diffusion Equation (2), and the electrical and optical equations must be solved self-consistently.

3 RESULTS

In broad-area DFB devices the longitudinal mode spectrum is symmetric with respect to the Bragg wavelength $\lambda_B=2\Lambda N_{eff}$, Λ being the grating period (for a first-order grating). The two modes whose frequencies are nearest to the stop-band reach threshold simultaneously [3-5], and single-mode operation can only be achieved by modifying the longitudinal structure, e.g. introducing a $\pi/2$ phase slip in the grating or a difference in reflectivity between the two facets. This happens with a single-stripe DFB laser as well [7]; taking the lateral carrier density profile into account results in a

change of the effective dielectric constant and therefore in a shift of the Bragg frequency. The spectrum as a whole remains substantially the same as for a broad-area model with the same longitudinal structure, it is only shifted following the Bragg frequency (Figure 1).

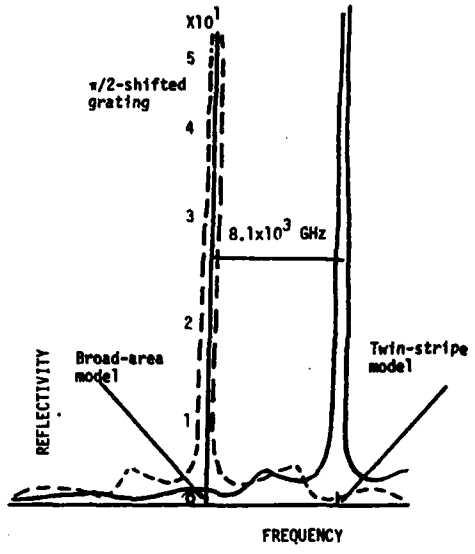


Figure 1: Broad-area (broken line) and twinstripe spectra with the same longitudinal structure

The shift in frequency is due to the "lateral" operator added to $\Delta\omega/v_g$ in (4): we can write the operator in brackets as

$$\frac{1}{v_g}\ \omega-(\omega_B - \frac{v_g}{2k_0N_{eff}}\ \frac{\partial^2}{\partial X^2} - \frac{k_0v_g}{2}\ \Delta\varepsilon)$$

$$= \frac{1}{v_g}\ [\omega-\omega'_B],$$

where ω'_B is the "perturbed" Bragg frequency and can be computed by use of the standard perturbation theory; clearly, the $\partial^2/\partial X^2$ term and the autiguiding term in (3) tend to increase the Bragg frequency, whereas the built-in index guiding tends to reduce it.

We have considered two different lateral structures: the first has two 3 μm wide stripes separated by a 3 μm gap, the second one is identical but with a built-in refractive index step equal to 10^{-3} localised in the gap; both structures have a $\pi/2$ phase slip in the grating for single-mode operation. When both stripes are at the same voltage, the carrier density and field profiles are symmetrical. Threshold is reached at 91.0 mA and 90.5 mA respectively, and the frequency of the main peak is higher by 8 GHz in the gain-guided structure than in the index-guided one, as the centre of the gap towards the high-pumped stripe, thereby reducing the effect of the index guiding. As to the antiguiding, we have expected (we are speaking now of $\nu=\omega/2\pi$).

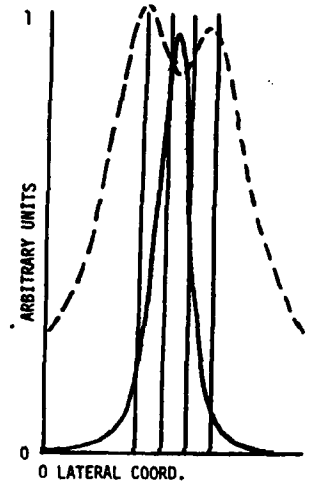


Figure 2: Carrier density (broken line) and field intensity (solid line) for asymmetrical pumping

Figure 2 shows the profiles when the stripe voltages are slightly different. The effect of the asymmetry is different in the two structures: for gain guiding, threshold is reached when the stripe voltages are 1.605V and 1.625V and the current is 91.05 mA, and the frequency of the main peak is shifted by -5.14 GHz (0.041 nm shift in wavelength); for index guiding, the voltages are 1.61V and 1.62V, the threshold current is 91.0 mA and the frequency is shifted by 8.6 GHz (0.068 nm). The carrier and field profiles are similar in the two structures, but in the index-guided one the field is less shifted.

The different signs of the frequency shifts for the two structures can be easily explained as follows. In the index-guided structure, the

index step is localised between the stripes: the asymmetric pumping shifts the field from to consider both the lateral shift of the field and the alteration of the carrier density profile, which decreases under the low-pumped stripe and (slightly) in the interstripe gap and increases under the high-pumped stripe. These are two competing effects, that must be considered together with the shift of the field: our calculations show that for small asymmetries this results in a decrease of the carrier-induced antiguiding. This is why the gain-guided structure experiences a decrease in frequency; when index guiding is present, its reduction is dominant and the frequency increases.

Of course, the situation is quite different when the asymmetry between the two stripes is larger: in Figure 3 we show the lateral

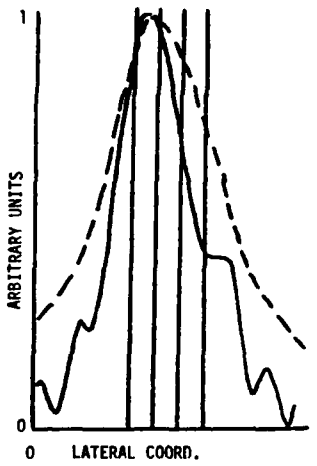


Figure 3: Carrier density (broken line) and field intensity (solid line) for very asymmetric pumping

carrier density and field profiles for the gain-guided structure, when the threshold current is 106 mA and the voltages are 1.55V and 1.74V; in this case, we have practically a (distorted) single-stripe laser. The frequency shift is as large as 2700 GHz (21 mm).

4 CONCLUSIONS

We have demonstrated that the longitudinal mode spectrum of a DFB laser can be "tuned" also by acting on its lateral carrier density profile, and that the frequency shifts obtainable in this way are comparable with those obtained by means of a direct "longitudinal" tuning: indeed, the tuning efficiency is of the order of 10 to 100 GHz/mA depending on the details of the lateral structure. Since ageing and degradation can alter the lateral geometry of a device, thereby introducing unwanted lateral asymmetries, the performance of a DFB laser can be significantly affected by the appearance of frequency and field shifts; the lateral structure must therefore be properly designed in order to minimise such effects or to provide for their compensation by means of some kind of lateral tuning mechanism.

REFERENCES

1 Dutta, N.A. et al, Appl. Phys. Lett., 48, (1986) 1501-1503

2 Murata, S. et al, Electron. Lett., 23 (1978) 12-14

3 Westbrook, L.D. et al, IEEE J. Quantum Eletr. 21 (1985) 512-518

4 Soda, H. and Imai, H., IEEE J. Quantum Eletr. 22 (1986) 637-641

5 Utaka, K. et al, IEEE J. Quantum Eletr. 22 (1986) 1042-1051

6 Syms, R.R.A., IEEE J. Quantum Eletr. 22 (1986) 411-418

7 Federighi, M., to be published. J. Appl. Phys. 62 (1987)

8 Kumar, T., Solid State Electron. 30 (1987) 21-31

9 Joyce, W.B., J. Appl. Phys. 53 (1982) 7235-7239

10 Joyce, W.B. and Dixon, R.W., Appl. Phys. Lett., 31 (1977) 354-356

11 McCall, S.L. and Platzman, P.M., IEEE J. Quantum Electr. 21 (1985) 1899-1904

12 Agrawal, G.P., J. Appl. Phys. 56 (1984) 3100-3109

IP.14. - INVITED PAPER

AlGaAs/GaAs MODULATION DOPED FETS FOR ULTRA HIGH SPEED SIGNAL PROCESSING APPLICATIONS

A. CHRISTOU

Research Center of Crete and Physics Department, University

of Crete, Iraklio, Crete, Greece.

MESFET and MODFET device technologies are reviewed in terms of device characteristics and processing difficulties. The MODFET discussion is then extended to include the planar self-aligned processes. The requirements for molecular beam epitaxy are presented. The results for ring oscillators are compared as well as a number of other device/circuit results.

INTRODUCTION

The development of a field effect transistor GaAs technology heralds a new generation of digital and analogue circuits for high performance systems such as telecommunications, high speed computers, advanced instrumentation and signal processing. Likewise the development of the MODFET (modulation doped field effect transistor) or HEMT also heralds a new generation of digital circuits due to its higher two-dimensional electron gas mobility and saturated drift velocity. These properties translate into a higher switching speed and current and lower parasitic capacitances for MESFET and silicon devices of comparable geometry. This paper will review the important aspects of AlGaAs/GaAs MODFETS with respect to signal processing applications. Specifically the following topics will be reviewed:

I. MESFETS versus MODFETS

II. The important aspects of the MODFET technology

a. Planar self-aligned process

b. Molecular beam epitaxy

III Device/Circuit Results and Summary.

In introducing a discussion of the MODFET technology one has to assess the material advantages of GaAs. As discrete components, GaAs FETs have been produced both as depletion mode and enhancement mode FETs. The enhancement mode FETs have resulted in a reduction of power consumption which allows one to increase their level of complexity up to LSI level.

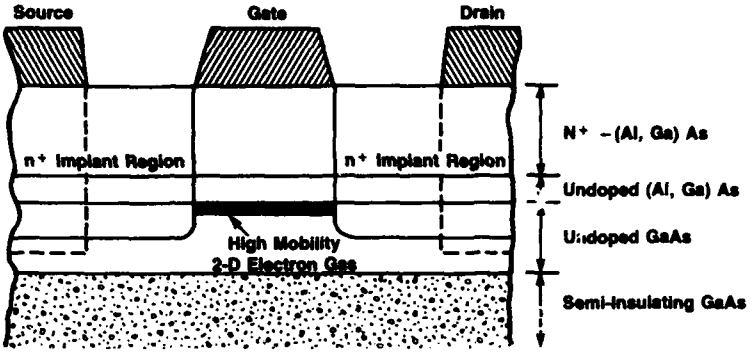
The basic FET structure for MESFETs and MODFETs remains the same: namely the recessed gate[1] and the self-aligned gate.[2] The conducting channel is formed in an n-type channel for the MESFET or a 2 dimensional electron gas channel for the MODFET. A Schottky barrier is then formed followed by the ohmic contacts. For high speed applications, the gate is recessed in order to minimize the parasitic resistance in series with the channel. In the self aligned structure, the gate itself is used to mask the formation of the source and drain. The most important difference between enhancement mode and depletion mode FETs is the thickness of the channel. Enhancement mode FETs require a thin channel which is fully depleted by the built-in potential of the Schottky gate. The depletion mode FETs are only depleted when the gate is biased negative with respect to the source.

I. MODFETs versus MESFETs

The cross-section construction of the MODFET is shown in Figure 1. Two characteristics of MODFETS are essential in order to achieve larger noise margins and high speed. These characteristics are the undoped GaAs channel and the large barrier height. The gate metallization makes contact through an n^+—AlGaAs

MODFETs FOR DIGITAL APPLICATIONS

- LARGE BARRIER HEIGHT
  - — Larger Noise Margins
  - — Higher Speeds
  - — High Yield MSI/LSI Circuits
- UNDOPED GaAs CHANNEL
  - — High Speed



Develop and Optimize MBE Materials Growth, Device Structures, and Fabrication Process to Achieve Ultra High Speed/Low Power MODFET Device/Circuits

Figure 1

Schematic showing MODFETS for a totally planar technology.

layer directly to an undoped AlGaAs spacer layer and an undoped GaAs region which contains the 2DEG layer. This configuration results in a large barrier height and therefore larger noise margins, higher speed and high yield circuits. The undoped GaAs channel due to the minimization of impurity scattering results in a higher transconductance and higher speed.

In order to compare MODFET and MESFET performance, ring oscillator data has been plotted for a number of high speed technologies as shown in Figure 2. At the present time E-Mode MODFETS have a gate delay time of 10-30 ps at a power dissipation per gate of 1 mW. In comparison E-Mode MESFETS show a gate delay time of 30-100 ps at 1-10 mW power dissipation.

The basic device physics of the MODFET may be understood through the examination of the conduction band diagram.[3,4] Figure 3 shows that a spatial separation of electrons from the ionized donors is possible at the n^+—AlGaAs/GaAs heterojunction. The conduction electrons are therefore present in the undoped GaAs potential well up to the Fermi level. The high ring oscillator delays are derived from the superior transport properties of the electrons in the 2DEG. The low resistance of the 2DEG permits the attainment of full device current at small voltages above threshold and also lowers the parasitic resistance of the MODFET. For a logic circuit, a low on-resistance is important to maintain a good noise margin.

II. MODFET/MESFET TECHNOLOGIES

The MODFET self aligned gate (SAG) processing technology is a 9 mask level process while a comparable E-MESFET SAG process is a 10 or 11 mask level process. The MODFET process is initiated by the substrate preparation followed by MBE growth, metal silicide, self-aligned implant, anneal, ohmic contact formation, isolation implant, first level metal, interlevel dielectric and via etch, second level metallization and bonding pad deposition. The E-MESFET SAG GaAs IC process includes a channel implant at the start of the process. The above process may include a submicron gate definition formed by electron beam lithography. The Schottky metallization for a SAG process is usually

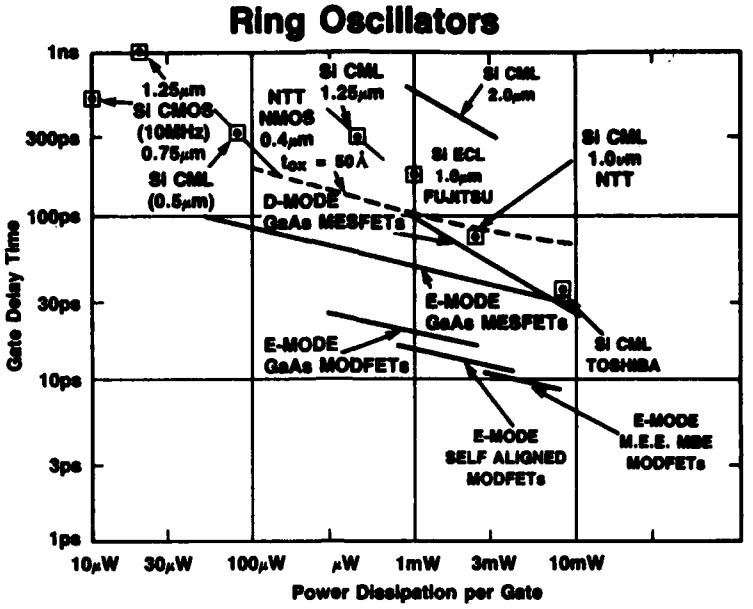


Figure 2

Gate delay and power dissipation data reported for ring oscillators.

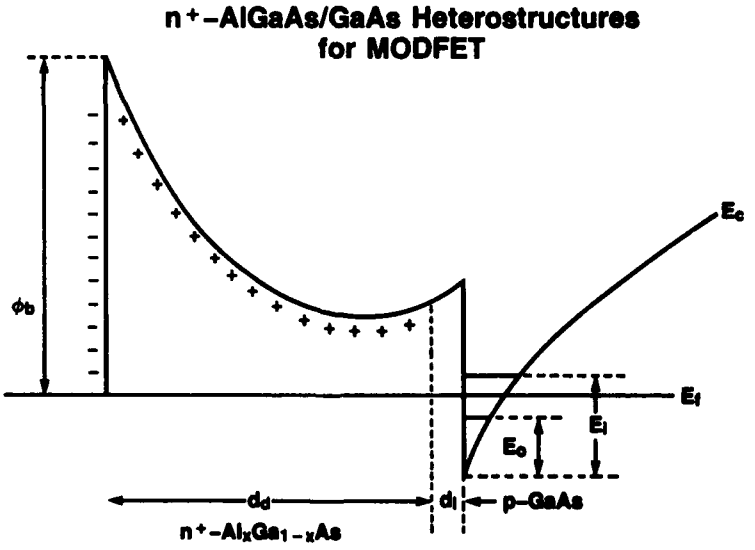


Figure 3

The energy band diagram for MODFETS

a TiW silicide, while metal 1 and 2 is TiAu. The interlevel dielectric is SiO_2 with upper and lower level being Si_3N_4.

The molecular beam epitaxy materials technology for achieving a high yield SAG process is shown in Table 1. The present technology utilizes 3 inch wafers and requires an electron mobility of 100,000-170,000 $cm^2/v.s$. Also achievable is the minimization of MBE surface defects to less than $100/cm^2$. Deep traps in the AlGaAs[5,6] may be eliminated through utilization of

TABLE I
MOLECULAR BEAM EPITAXY MATERIALS STATUS

- **3 INCH GaAs PROCESS**
- **EXTREMELY HIGH ELECTRON MOBILITY**
  — 77K MOBILITY OF MODFETs: 100,000 — 170,000 cm^2/Vs
- **LOW MBE SURFACE DEFECTS**
  — KILLER OVAL DEFECTS < $100/cm^2$
- **DEEP TRAPS ELIMINATED WITH SUPERLATTICES**
  — DX CENTERS REDUCED BY FACTOR OF 30
- **MATERIAL CONTROL**

Parameter	Reproducibility	Uniformity
Thickness	3%	2% in central 2.5 inch of 3 inch wafers
Doping	3%	2% over 3 inch
Aluminum Composition	2%	< 0.5% over 3 inch

- **V_T UNIFORMITY OF MODFETs: = ±15mV OVER CENTRAL 2 INCH**

superlattices. The trapping problem in the MODFET is the most serious problem affecting its large-scale logic application. The main trapping center is associated with donors in the AlGaAs. At low temperatures the centers may be ionized by light and free-carriers will remain in the AlGaAs and with recapture will lead to persistent photoconductivity.

Utilization of a superlattice has been able to eliminate many of the trapping effects and has resulted in a superior threshold voltage uniformity. This is shown in Figure 4 for a gate array and indicates only a standard deviation of 0.0084 volts from an average threshold voltage of 0.18V for an enhancement mode MODFET. The drain-source voltage-drain current characteristics are shown in Figure 5 for a 1.0 μm MODFET and indicates a g_m of approximately 225 mS/mm at room temperature. A source resistance of 0.7 ohm-mm was also measured for the MODFET. The g_m at 77 K increased to 275-400 ms/mm. The improved noise margin performance is shown in Figure 6 where the noise margin has been measured to be 390 mV.

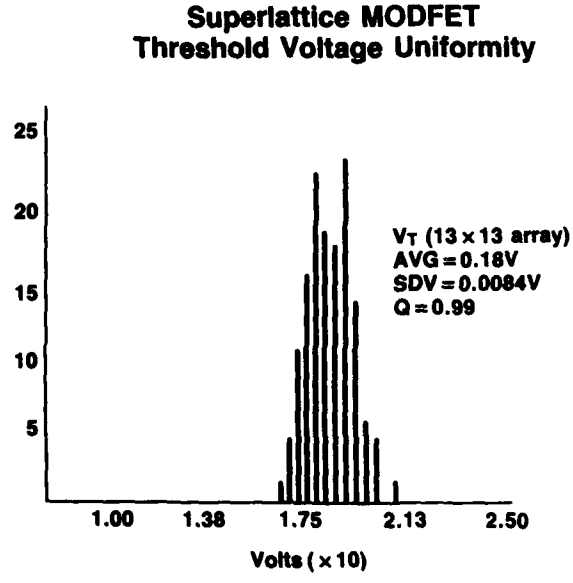


Figure 4

Superlattice threshold voltage uniformity for MODFETS.

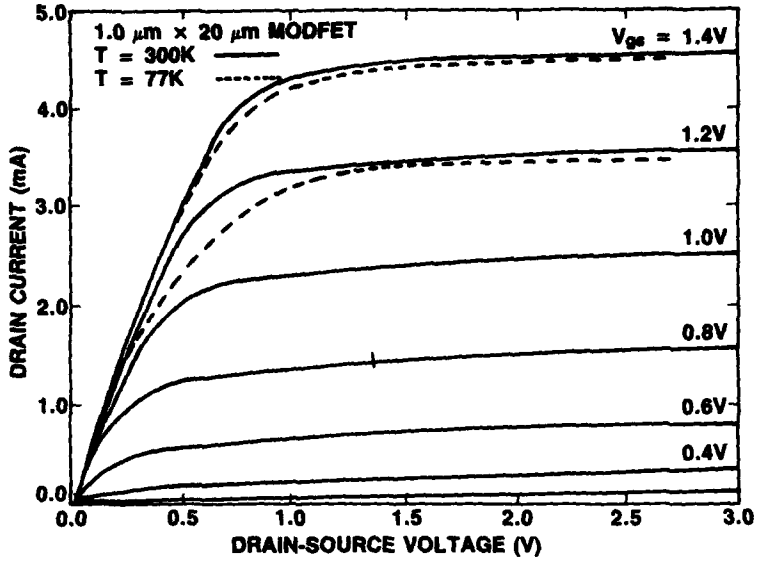


Figure 5

IV characteristics for a SAG MODFET.

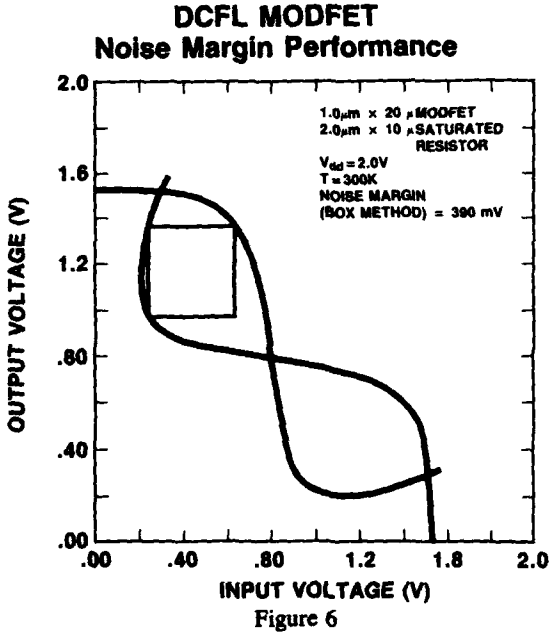


Figure 6

MODFET noise margin performance.

III. SUPERLATTICE MODFET I.C.s

The n^+AlGaAs layer of the MODFET may be replaced by a superlattice layer in order to have a n^+ (Si)-GaAs : i-AlAs superlattice charge control mechanism. The Si is now only within the GaAs resulting in a higher silicon activation, and a larger charge density. As a result of the superlattice, elimination of persistent photoconductivity is possible, as well as reduced threshold voltage shifts. The above configuration also eliminates the presence of DX centers since only the GaAs is doped. Figure 7 shows the band diagram of a superlattice MODFET showing the undoped Al_xGa_{1-x} As and the silicon doped GaAs. The spacer region is typically 50-100 Å thick followed by an undoped GaAs buffer. The I_{DS} versus V_{DS} characteristics show no evidence of trapping and in addition the 77 K characteristics are very similar as the 300 K results as shown in Figure 8. The DLTS spectra (deep level transient spectroscopy) for the n^+ GaAs /i-AlGaAs superlattice shows a 30-fold reduction in the trap concentration of over conventional MODFET structures. In addition, the threshold voltage variation has been decreased significantly.

A comparison of the propagation delay and power dissipation for gate arrays and MODFETs is shown in Figure 9 indicating improved performance for the circuits with the superlattice charge separation region. Both MESFET and MODFET technologies are

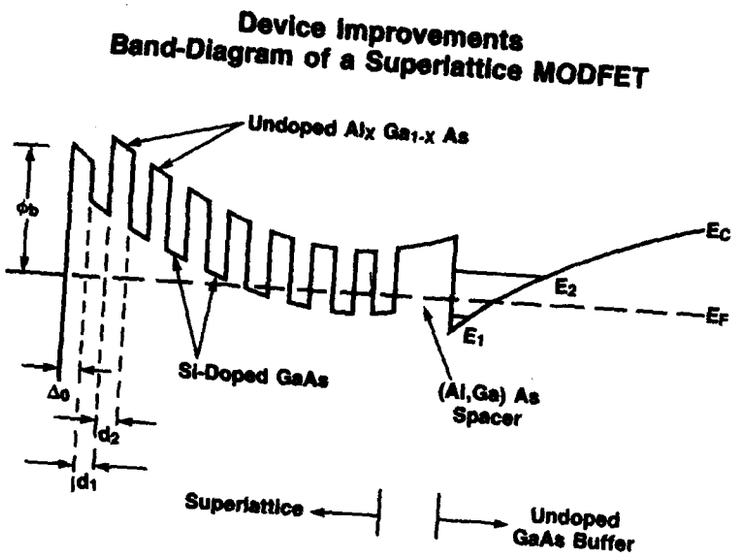


Figure 7

The band diagram of a superlattice MODFET.

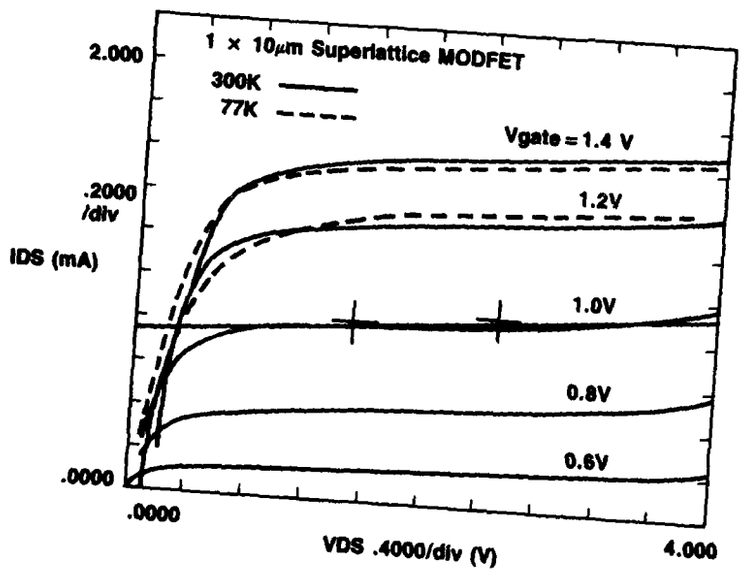


Figure 8

IV characteristics of a superlattice MODFET.

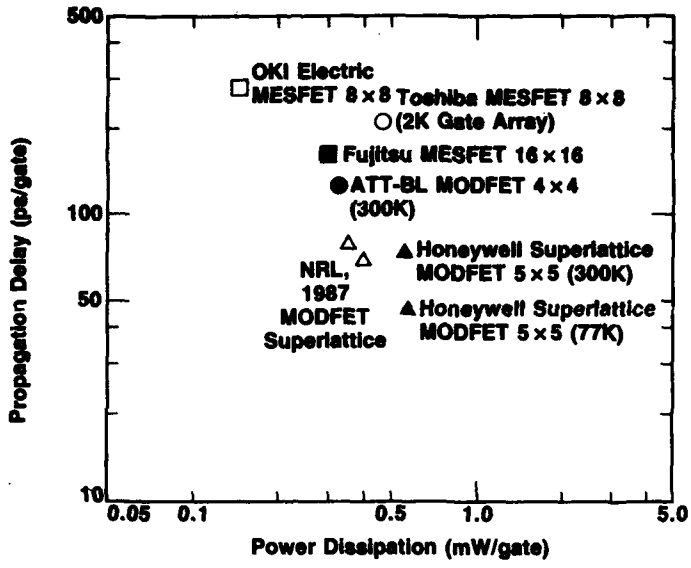


Figure 9

Power dissipation data for MODFETS and MESFETS.

compared. The optimum result reported to date is a 50 ps/gate delay time at a power dissipation of 0.5 mW/gate.

IV. CIRCUIT RESULTS AND CONCLUSIONS

Results from a self aligned gate MODFET triple cell for both and/nor and nor gate delay has resulted in typical gate delay times of 40-75 ps/gate and 0.5-2.0 GHz clock frequency. Power dissipation for the MODFET self aligned gates is of the order of 0.5-2.0 mW/gate. At elevated temperatures in the Banville Triple Cell configuration optimum clock frequency is typically observed to decrease. Table II summarizes the MODFET results reported for digital electronics. The circuit of greatest complexity is a 5 × 5 multiplier with a gate delay of 72 ps measured at 77 K. The ring oscillator results at 2.33 GHz indicate a gate delay of 8.5 ps at 77 K. The 64 bit SRAM circuit showed the least power consumption of 0.27 mW/gate.

TABLE II

AlGaAs/GaAs MODFETs FOR DIGITAL ELECTRONICS

CIRCUITS	GATE COUNT	CIRCUIT SPEED	GATE DELAY	POWER	COMMENTS
Ring Oscillator	25	1.75 GHz	11.6 ps	1.55mW/gate	300K FO = 1
		2.33 GHz	8.5 ps	2.59	77K FO = 1
Divide-by-4	25	4.3 GHz	46.5 ps	2.0	300K FO = 2
64 bit SRAM	250	1.1 nsec		0.27	
5 × 5 Multiplier	370	1.8 nsec	72 ps	0.43	300K FO = 2.2
		1.1 nsec	44 ps	0.73	77K
Banville	233	1.7 GHz	42 ps		300K FO = 2.2 fully functional -55 to 200C

In conclusion, the present MODFET technology based on a planar self-aligned process is viable for ultra high speed signal processing applications. This technology offers the advantage of larger noise margins and higher speeds.

REFERENCES

[1]. E. Takeda, H. Kume, T. Toyabe, and S. Asai, *IEEE Trans. Electron Devices*,, Vol. ED-29, p. 611, 1982.

[2]. F.A. Buot, and K.J. Sleger, Solid State Electronics 27, 1067 (1984).

[3]. N.T. Linh, in Applications of GaAs MESFETs, R. Soares, J. Graffeuit, and J. Obregon, (Artech House 1983) Ch. 9.

[4]. R. Mimue, A. Hiymizu, T. Fujii, and K. Nanbu, Kpn. J. Appl. Phys. 19, L225 (1980).

5. R. Fischer, T.J. Drummond, et. al. IEEE Trans. Electron Dev. ED-31, 1028 (1985).

[6]. A. Kastalsky, and R.A. Kiehl, IEEE Trans. Electron. Dev. ED-31, 414 (1986).

Session A4.3

Bipolar Technology

Chairman: A. Wieder

Thursday, September 17, 1987

DOUBLE SELF ALIGNED BIPOLAR TRANSISTORS USING SALICIDE CONTACTS

H. Kabza, M. Reisch, V. Probst, W. Böhm, J. Fertsch, H. Schaber and H. Eggers*

Siemens AG, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, FRG
*Siemens AG, Components Division, Balanstr. 73, D-8000 München 80, FRG

A new double self aligned fabrication process for bipolar transistors is presented by applying the self aligned silicide (salicide) process using Pt to poly-Si self aligned devices Very low sheet resistivities for the external base region and therefore low base resistances are obtained. Temperature stability of the devices for annealing at 700°C for 60 min. is demonstrated. No silicide bridging is observed.

1. INTRODUCTION

Recent improvements in bipolar technology and circuit performance are mainly exploiting the benefits of polysilicon self aligned- (PSA-) processes [1,2]. The key feature of these processes is self alignment between emitter and extrinsic base, both of these regions being contacted by highly doped polysilicon layers. This concept leads to a drastic reduction in base-collector junction area and thus C_{bc}. Furthermore, extrinsic base resistance R_{bext} is reduced significantly due to the self alignment of the annular p^+-polysilicon base contact layer to the emitter region.

This is no longer sufficient, however, with further reduction of emitter widths down to the submicron range, especially when elongated emitter stripes are used for minimizing the internal base resistance R_{bint}. This is because R_{bint} is still limited by the p^+-poly-Si sheet resistance (usually 100 Ω/□ or higher) and thus will again dominate the total base resistance R_b and limit device performance. Furthermore, unfavourable asymmetric current distributions in the active region may occur due to the voltage drop within the p^+-polysilicon layer [3].

2. PROCESS FLOW

These problems may be resolved by introducing a second self aligned process step, namely the formation of a self aligned silicide (salicide) layer on top of both p^+ and n^+ polysilicon layers. The process flow first follows conventional bipolar processing using a buried n^+ collector region, n^- epitaxy and recessed oxide isolation (fig. 1a). Next a sandwich of p^+ poly-Si and SiO_2 is deposited by LPCVD and patterned with vertical sidewalls using reactive ion etching (RIE). By depositing another CVD oxide layer and applying an

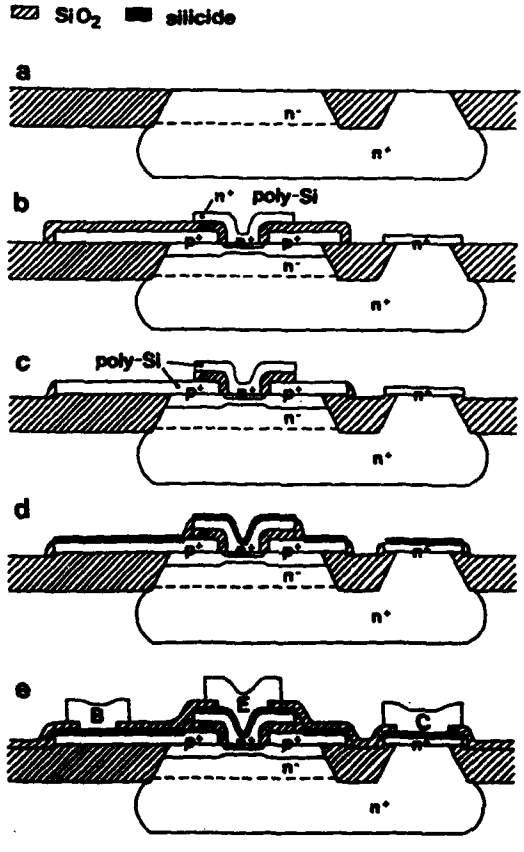


Fig. 1: Process flow for forming salicide contacted double self aligned bipolar transistors.

unmasked RIE etch-back, sidewall spacers are formed at the p^+ poly-Si electrodes. Base implant, n^+ poly-Si deposition, emitter drive-in and patterning of the n^+ poly-Si layer yield the self-aligned emitter base structure depicted in fig. 1b.

At this point the familiar PSA process flow [1] is left. The n^+ poly-Si pattern is used as a mask for another oxide RIE step, thus exposing the base (p^+) polysilicon layer except where covered by the emitter (n^+)-polysilicon (fig. 1c). A second oxide spacer is then formed by depositing 200 nm of SiO_2 (LPCVD-TEOS) and back-etching this layer in an unmasked RIE step. This second spacer serves to avoid silicide bridging at the edge of the n^+-polysilicon layer. In our experiments platinum was then sputter-deposited to a thickness of 35 nm. Silicidation was performed at 370°C in a wet oxygen ambient; unreacted Pt was subsequently etched off in aqua regia. The resulting structure (fig. 1d) shows complete silicidation of all polysilicon electrodes except for the base polysilicon in the narrow n^+/p^+ poly-Si overlap region. The silicided part of the base-polysilicon forms a ring of low sheet resistance around the emitter region. Oxide deposition, definition of contact holes and standard TiW/AlSiTi metallization complete the process (fig. 1e). A SEM picture of the emitter region of a transistor before opening the contact holes is shown in fig. 2.



Fig. 2: SEM-cross section of the emitter and extrinsic base region of a Pt-salicided transistor before the opening of contact holes.

3. BASE RESISTANCE AND DEVICE PERFORMANCE

The sheet resistance of poly-Si in our standard process is typically 150 Ω/□ . Using the Pt-salicide process outlined above it is lowered to 7 Ω/□ . In fig. 3 a calculation [3] of the total base resistance and its external component (emitter area 2x8 μm^2) with and without salicide is shown as a function of the collector current density.

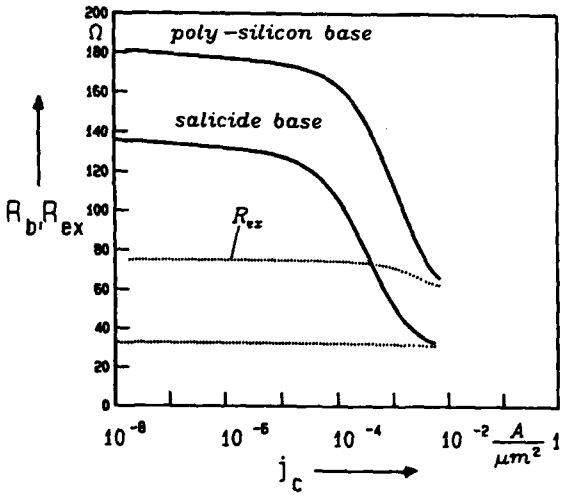


Fig. 3: Total base resistance and its external component of a transistor with and without Pt-salicide as a function of collector current density.

The external component is reduced from 75 Ω to 33 Ω and is mainly determined by the sheet resistance (appr. 50 Ω/□) of the non-silicided part of the extrinsic base region (overlap area between p^+- and n^+-poly-Si).

The impact on performance is demonstrated by the gate delay times of push-pull ECL ring-oscillators: a sensitivity analysis based on SPICE and the appropriate base resistance data gives a reduction of the gate delay time of typically 20 % depending on emitter geometry.

4. STATIC DEVICE PROPERTIES AND THERMAL STABILITY

Crucial issues for the application of the salicide process in the production of high speed bipolar circuits are device properties, stability with respect to thermal stress after silicidation and the yield obtainable.

To clarify the impact of the new base contact process on the stability of our transistors with respect to thermal stress the devices were annealed after silicidation. Fig. 4 shows input and transfer characteristics as well as the current gain vs. Ic for a device annealed for 1 h at 700°C after silicidation.

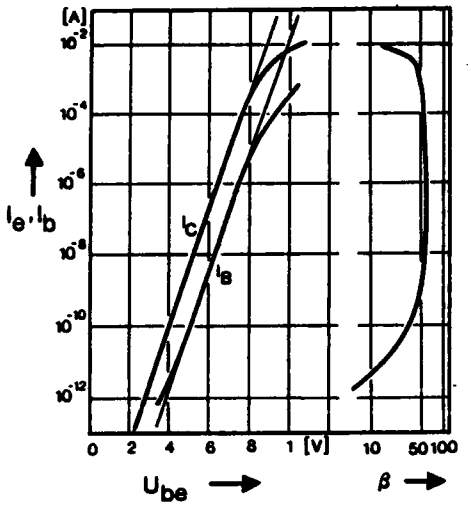


Fig. 4: Semilogarithmic plot of I_C, I_B vs. U_{BE} and double logarithmic plot of I_C vs. β of a sample with an additional annealing cycle (700°C, 60') after silicidation.

A comparison of the reverse bias characteristics of emitter base diodes with and without salicided contacts (fig. 5) clearly shows that no silicide bridging occurs. Both devices show comparable tunneling characteristics which are determined by the doping in the sidewall diode [4]. Differences in the reverse bias characteristics of polysilicon and salicide contacted transistors were observed only at pA current levels. We investigated the temperature dependence of the additional current component and found the slope of the corresponding activation energy plot (fig. 6) to be 540 meV ≈ $E_G/2$ in close accordance to what is to be expected from the SRH deep trap model [5].

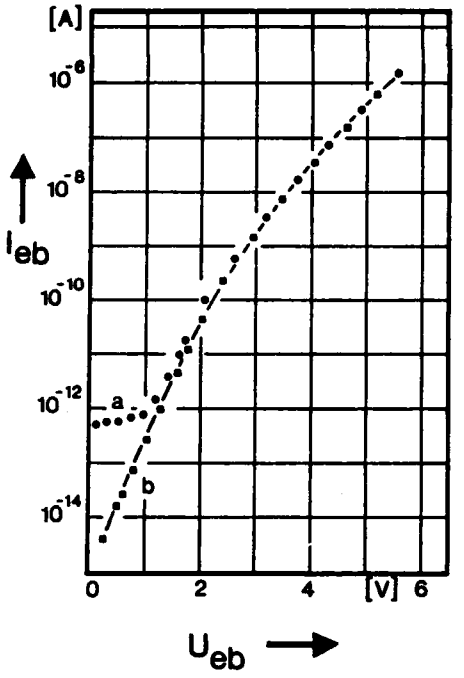


Fig. 5: Comparison of reverse bias characteristics of polysilicon (b) and salicide contacted (a) transistors.

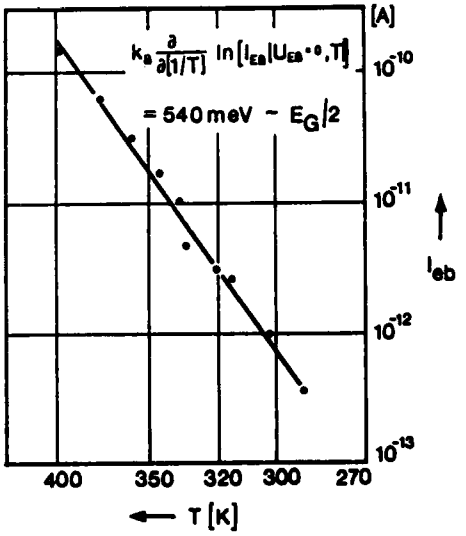


Fig. 6: Activation energy plot of the additional current component observed in the EB diode.

Since generation currents of this order of magnitude were found also in batches without salicided base contacts we further investigated the generation current per unit volume in the reverse biased base collector junction and found

$$\frac{\delta j}{\delta V} < 10^{-12}\ A/\mu m^3 \qquad (4.1)$$

There is no evidence that this additional generation current component is caused by Pt impurities but we may conclude that (4.1) provides an upper limit for the increase in generation current to be expected.

5. CONCLUSIONS

We conclude that the salicide process presented allows the production of BJTs with nearly ideal characteristics even after considerable thermal stress - a fact especially important for a subsequent multilayer metallization process.

The sheet resistance of the external base was lowered to 7 Ω/□ and so the total external base resistance could be reduced drastically.

As silicide bridging was not observed this is obviously not a yield limiting factor.

Of course the process flow described above may as well be used with other silicides allowing self aligned silicidation like $TiSi_2$ or $CoSi_2$. For these thermally very stable materials, favourable combinations of emitter drive-in and silicidation steps might be envisaged.

REFERENCES

[1] T.H. Ning et al., IEEE Trans. on Electron Devices ED 28, 1010 (1981)

[2] A.W. Wieder, Siemens Forsch. u. Entw.-Ber. 13, 246 (1984)

[3] J. Fertsch, H. Klose and W.R. Böhm, Proc. of ESSDERC 1986, p. 65

[4] H. Schaber, L. Treitinger and A.W. Wieder Proc. of ESSDERC 1984, p. 342

[5] Sze, S.M., Physics of Semiconductor Devices, p. 90 (Wiley, New York, 1981)

Speed-Power Relation of Modern Bipolar Technology

P. Weger

Siemens AG, Central Research and Development, Munich, FRG

H.-M. Rein

Ruhr-University Bochum, Institute of Electronics, FRG

The recent progress in silicon bipolar circuits shows a remarkable leap to higher speed and lower power consumption. This fact results mainly from two achievements: self-alignment and polysilicon emitter contact. In this work, the high-speed and low-power capability of a modern bipolar technology is demonstrated. It is shown that even by use of rather conservative 2μm lithography, remarkable results can be achieved.

First, the advantages of the polysilicon self-aligned technology (PSA) are shown by comparing this technology with a standard (buried collector) technology (SBC). The simplified cross-sections of the corresponding transistors are given in fig. 1.

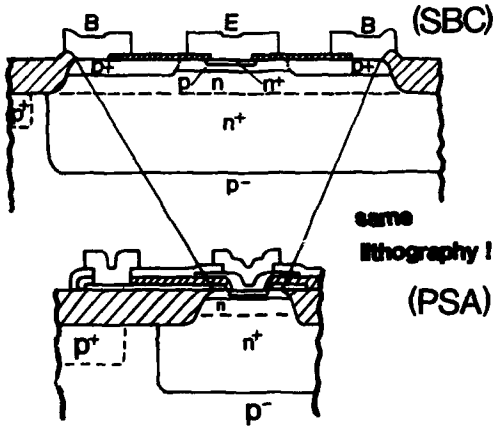


Fig.1 Cross-sections of transistors for standard (SBC) and PSA technology.

The PSA technology developed for the production of complex gate arrays was described elsewhere [1]. It is a double-polysilicon technique with oxide spacer between emitter contact and extrinsic base contact (see fig.2). The polysilicon layers are acting both as diffusion sources as well as contact materials. From fig.1 it can clearly be seen that the extension of the external base region of the PSA is drastically reduced compared to SBC. To a large extent, it is replaced by the polysilicon layer on top of a thick oxide. These features result in superior transistor parameters compared to the SBC technology. In table 1, transistor parameters of both technologies are given on the basis of an emitter with a length of 20μm and a 2μm lithography. The main differences can be noticed in the base spreading resistance R_b, especially in the external resistance R_{bex}, the collector-base capacitance C_{jc}, and the forward transit time τ_F. Due to the oxide spacer,



Fig.2 REM photograph of the cross-section of a PSA transistor

Table 1 Comparison of the transistor parameters for PSA and SBC technology. Emitter mask size 2x20μm^2

	PSA	SBC
Effective emitter size	1.4 x 19.4 μm^2	2 x 20 μm^2
Base resistance $R_b = R_{bi} + R_{bex}$	60 Ω + 8 Ω	87Ω + 46Ω
Emitter-base capacitance C_{je0}	114 fF	196 fF
Collector-base capacitance C_{jc0}	70 fF	181 fF
Collector-substrate cap. C_{js0}	140 fF	304 fF
Base transit time τ_{F0}	8 ps	13 ps

the effective emitter width is $1.4\mu m$ for the mask size of $2\mu m$. Therefore, the internal base resistance R_{bi} is about 30% smaller in the case of self-aligned emitter-base configuration at the same sheet resistance of the internal base (about $10k\Omega/\square$).

The use of polysilicon as a source for the diffusion of arsenic enables shallower emitters and smaller base widths leading to smaller transit time τ_F. Moreover, a further reduction of transistor area at low R_b can be achieved by use of silicide or "polycide" instead of polysilicon base contact [3].

The capability of technologies for producing high-speed ICs can be well demonstrated with circuits like static frequency dividers. Therefore, 8:1 dividers based on master/slave D flip-flops with the inverted output fed back to the data input have been realized for this purpose.

First, the circuit was designed and optimized for maximum operating frequency. Optimization included the resistances, currents, and voltage swings as well as the geometry of all individual transistors (cf. [4]-[6]). The circuit diagram and further details are described in [7], the photograph of the chip is shown in fig. 3 with a chip area of $0.9mm^2$ including bond pads. At a total



Fig.3 Chip photo of the 8*GHz* frequency divider

power consumption of $430mW$ for the 8:1 frequency divider including the output buffer with 50Ω load, a maximum input frequency f_{max} of 8*GHz* was achieved. This is the highest frequency for static frequency dividers based on a $2\mu m$ technology ever reported to the authors' knowledge. The input and output signal at this frequency is shown in fig. 4.

Fig.5 presents the minimum required input voltage vs. input frequency. It can be seen that at frequencies somewhat lower than 8*GHz*, the input voltage can considerably be reduced.

For comparison, with a $2\mu m$ standard technology, similar to that in fig. 1, a maximum input frequency of 3.4*GHz* has been achieved [8].



Fig.4 Input and output signals of the 8:1 frequency divider at 8*GHz* input frequency

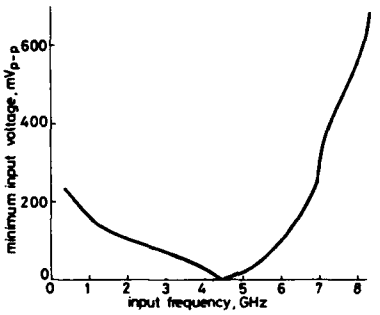


Fig.5 Input sensitivity of the 8*GHz* divider

These results can be well compared because both circuits have been designed using the same methods and, moreover, consume about the same power. This demonstrates a speed advantage of the PSA technology by more than a factor of two for nearly the same lithography. This result has been confirmed by investigation of other types of high-speed circuits on the basis of both technologies [8].

The question is how much f_{max} is reduced if the power consumption P is lowered. For the circuit principle used, f_{max} is about inversely proportional to the gate delay t_d. Assuming that all transistors have the same size and that the relation (2) given in [4] holds in rough approximation also for the present circuit principle, we get

$$\frac{1}{f_{max}} \sim t_d \approx (1+k_1\frac{R_b}{R})\cdot(k_2\tau_F+k_3RC_{je})+k_4RC_{eff}. \quad (1)$$

This expression should only indicate the basic relationships. At fixed voltage swings and resistance ratios, the factors $k_1 \cdots k_4$ are constant,

C_{eff} is an effective capacitance representing the influence of junction and wiring capacitances (cf. [4]). R is an arbitrary reference resistance (e.g. the load resistance of the flipflops) which changes with varying power consumption P as

$$R \sim \frac{1}{P} \quad . \tag{2}$$

At small power consumption, the last term in (1) is dominating.

As already demonstrated in [4], [5], the reduction of f_{max} with decreasing P (i.e. rising R) can be kept small by reducing the length of emitters, l_E, proportionally to P. In this way, the terms $\frac{R_b}{R}$ and RC_{jc} in (1) remain nearly constant. Furthermore, RC_{eff} is only comparatively slightly enhanced as long as the lengths of emitters are sufficiently large and the junction capacitances are dominating over the wiring capacitances.

This adaption of transistor size to power consumption can only be obtained by changing of the masks. Therefore, first the dimensions of the transistors have been maintained and the power consumption has been reduced by increasing only the sheet resistance of the polysilicon resistors. The experimental results obtained in this way are shown in fig. 6. At $P \approx 60mW$ still $4GHz$ are achieved.

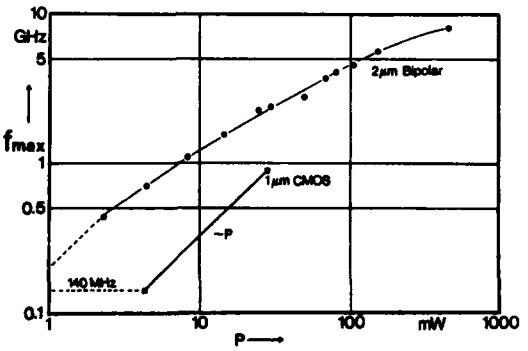


Fig.6 Maximal input frequency of the frequency dividers vs. power consumption

For further reduced power consumption, a second circuit with transistors all of minimum size and only with one base contact has been realized (emitter mask dimensions $2 \times 4\mu m^2$). The maximum operating frequency of this circuit is $3.8GHz$. Below about $2.5GHz$ ($P<30mW$), it consumes less power than the first circuit designed for maximum speed. At $P=10mW$, still more than $1GHz$ is achieved.

It should be emphasized that the values of f_{max} at medium and higher power consumption would be considerably higher than those shown in fig. 6 if the dimensions of transistors were adapted to P. At small power consumption and minimum transistor dimensions, higher values of f_{max} are achievable if the resistivity of the epitaxial collector ($\rho_{epi} \approx 0.3\Omega cm$) is enlarged resulting in a reduction of C_{jc}.

The results of the present work confirm that not only the maximum achievable operating speed is more than two times higher with a modern PSA technology than with a standard SBC technology using about the same lithography, but also prove the PSA technology to be very well suited for high-speed circuits with low power consumption.

Additionally, it should be mentioned that the ideal low-current characteristics of the PSA transistors [1] enable us to realize circuits with extremely low power consumption together with high-speed circuits on a single chip. Therefore, for many applications the PSA technology is a serious rival for modern CMOS technologies in the low-power regime.

For comparison, fig. 6 contains the results $f_{max}(P)$ of a 8:1 frequency divider realized in a $1\mu m$ CMOS technology. At fixed $f_{max}(\leq 1GHz)$, the power consumption of the CMOS circuit is drastically higher than that of the bipolar circuit. Using the same lithography for both technologies, the difference is much more enhanced. These results are mainly a consequence of the much smaller voltage swings of the bipolar circuits.

However, it should be pointed out that this comparison looks better for the CMOS technology, if -as in many logic circuits- the medium switching period of the gates and flipflops is large compared to the gate delay.

References:

[1] A. W. Wieder, "Self-aligned bipolar technology -new chances for very high-speed digital integrated circuits", Siemens Forsch.-& Entwicklungsber., Vol. 13, pp. 246-252, (1984)

[2] T.H. Ning and D.D. Tang, "Bipolar trends", Proc. of the IEEE, Vol. 74, pp.1669-77, (1986)

[3] H. Kabza, M. Reisch, V. Probst, W. Böhm, J. Fertsch, H. Schaber and H. Eggers, "Double self-aligned bipolar transistors using salicide contacts", this Conference A4.3.1

[4] R. Ranfft and H.-M. Rein, "A simple optimization procedure for bipolar subnanosecond ICs with low power consumption", Microelectronics J., Vol. 13, No. 4,pp. 23-28, (1982)

[5] R. Ranfft and H.-M. Rein, "High-speed bipolar circuits with low power consumption for LSI - a comparison", IEEE J. Solid-State Circ., Vol. SC-17, pp. 703-712, (1982)

[6] R. Reimann and H.-M. Rein, "A 4:1 time divison multiplexer IC for bit rates up to 6*Gbit/s* based on a standard bipolar technology", IEEE J. Solid-State Circ., Vol. SC-21, pp. 785-789, (1986)

[7] P. Weger, L. Treitinger, R.Reimann, H.-M. Rein, "Static 7*GHz* frequency divider IC based on a 2μm Si bipolar technology", Electronics Letters, Vol. 23, pp. 192-93, (1987)

[8] H.-M. Rein, "Gigabit/s silicon bipolar ICs for future optical-fiber transmission systems", ESSCIRC 1987 (invited paper).

Study on Current Transport Mechanism in Amorphous SiC Emitter HBT

Kimihiro Sasaki and Seijiro Furukawa

Department of Applied Electronics,
Graduate school of Sci. & Eng., Tokyo Inst. of Tech.,
4259 Nagatsuda, Midori-ku, Yokohama 227, Japan

Current transport model and current gain degradation mechanisms of a silicon heterojunction bipolar transistor with amorphous SiC emitter are discussed by observing various characteristics of amorphous SiC films as the emitter material. Under such study current gain is remarkably enhanced by using low resistance micro-crystalline Si, and current gain as high as 480 has been obtained.

1. INTRODUCTION

We have been investigating Si heterojunction bipolar transistors (HBTs) with amorphous SiC (a-SiC) wide band gap hetero-emitter [1], [2] in order to realize future high speed bipolar transistor and its system. The device has several prospects:

1) High current gain can be preserved due to wide band gap emitter even under heavily doped base condition, which means reduction of base resistance.
2) In order to fabricate the very thin and heavily doped base region, the process temperature must be low. The a-SiC emitter can overcome this limitation, because the deposition temperature of a-SiC is so low as to prevent redistributi of base impurities.
3) Moreover, It is considered that this emitter has good process compatibility with conventional bipolar transistors having poly Si emitter, since the poly Si emitter can be easily replaced by this emitter.

In this paper, at first we will explain fabrication process of a device and fundamental properties of a-SiC films. Then the cause of lower current gain than the value expected theoretically from the band structure will be cleared by considering a current transport mechanism on the a-SiC/c-Si heterojunction. Furthermore, after indicating means such as useful micro-crystalline Si to enhance current gain. Actually, we will show an example achieving current gain as high as 480.

2. DEVICE FABRICATION PROCESS

Conventional Si process was used for fabrication of the HBTs.

A field oxide film was grown thermally on a chemically precleaned 7 Ωcm (111) n/n^+ Si substrate and a window was cut into it. Ion implantation (B^+, 8×10^{12} cm^{-2}, 80 keV) was carried out for fabricating base region. A second oxidation(1025 °C, 30 min), base drive-in diffusion and activation anneal were carried out simultaneously. An emitter window was cut into the oxide, the base surface was cleaned by RCA boiling again. Immediately after the sample was dipped in the HF solution in order to remove a chemical oxide, the sample was loaded into the discharge chamber of plasma CVD system wherein an a-SiC film was deposited. Then, the a-SiC film was etched by CF_4 RIE etching except emitter region. After finally Al electrodes were fabricated, treatment such as thermal or H_2 anneal was not applied. A cross-section of the transistor is shown in Fig.1.

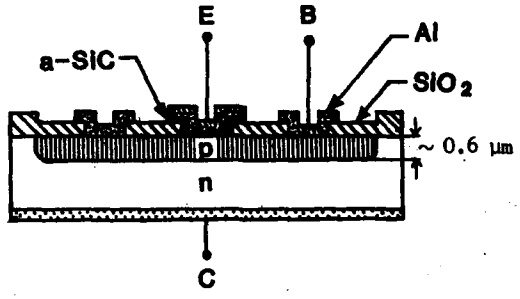


Fig.1. Cross-section of transistor.

Depo. Method	:L-coupled plasma CVD
Sub. Temp.	:350 ℃
Dischage Power	:60 W
Pressure	:0.4 Torr
Usig Gases	:SiH_4 5% (He, H_2 dilution)
	CH_4 100%
	PH_3 1% (He, H_2 dilution)
Depo. rate	: 40 Å/min

Table 1. Deposition conditions of a-SiC.

3. RESULTS AND DISCUSSION

3.1 Charactarization of a-SiC Films

a-SiC films were deposited by using inductive coupled plasma CVD system. Deposition conditions are shown in table 1.

Carbon content X, conductivity σd and optical band gap $Eg_{(opt)}$ were measured by EPMA, I-V characteristics between two electrodes and photo absorption respectively for H_2 and He diluted gas sources. In those measurements, suitable substrates (metal for carbon content, Corning 7059 glass for conductivity and optical band gap) were chosen. The results were shown in Fig.2 as a parameter of CH_4 fraction.

As the CH_4 fraction increases, X and $Eg_{(opt)}$ monotonously increase, however, σd rapidly decreases. C content in the a-SiC film prepared by H_2 dilution gases is about half of C content of the film by He dilution. This can be explained that decomposition of CH_4 in H_2 dilution case would not be so remarkably occurred as He dilution case. The conductivities of those two films at the same C content are almost same. Therefore, it is concluded that the conductivity principally depends on the C content in the film, rather than on the kind of dilution gases.

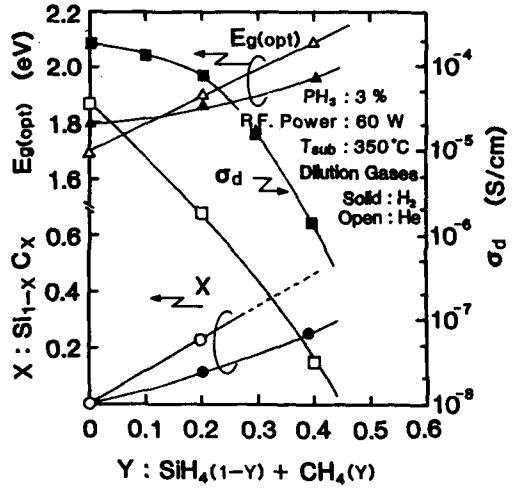


Fig.2. Fundamental properties of a-SiC films prepared from He or H_2 dilution gases.

3.2 Current Gain Degradation Mechanism

Typical V_{CE}-I_C characteristics for a-SiC emitter HBT are shown in Fig.3. A current gain of 180 is achieved in low collector current. This value corresponds to emitter Gummel number of 2×10^{14} s/cm^4, suggesting the wide band gap emitter effect [2].

A factor determining the current gain was indicated by using a band structure, resulting that we could explain large difference of current gain between pnp type HBT and npn type HBT [1]. However current gains obtained were not always high comparing with one theoretically predicted from the band structure.

In order to clarify mechanisms of such relatively low current gain, the characteristics of emitter current and current gains were measured as the function of V_{BE}, and are shown

in Fig.4. It is observed that the current gain degradation starts at the same point at which emitter current starts saturate. This behavior could be interpreted by a current transport model at an a-SiC/c-Si hetero-junction as shown in Fig.5.

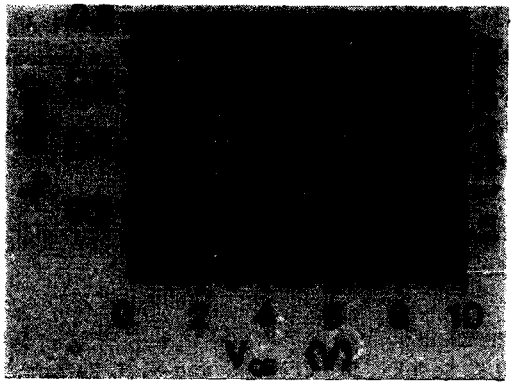


Fig.3. Curve-tracer presentation of V_{CE}-I_C characteristics for the a-SiC hetero-emitter Si-HBT.

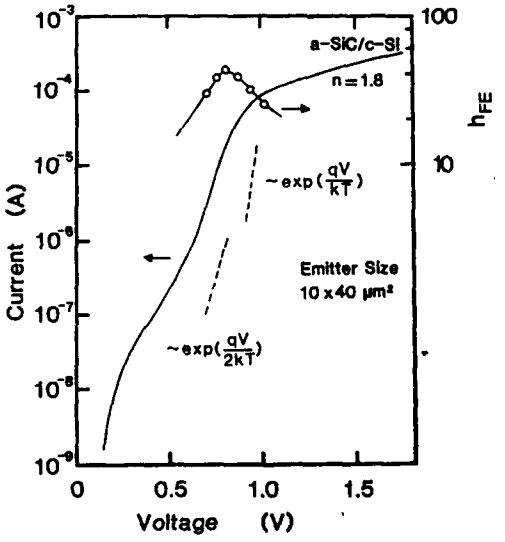


Fig.4. Emitter-base diode current and current gain dependences on V_{BE}.

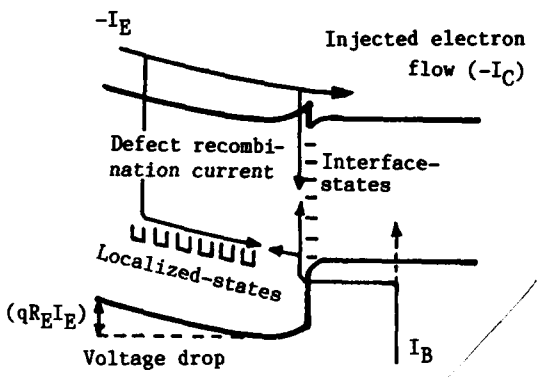


Fig. 5. Current transport model for a-SiC/c-Si hetero-structure.

It is considered that since the a-SiC film shows high resistivity as shown in Fig.2, carrier concentration in a-SiC is extremely small. Therefore, a current flowing through the emitter would be in the drift-current-limited even at such relatively small emitter current region. On the other hand, the interface recombination current and the defect-assisted current flowing via localized states induced by the defects in a-SiC which contribute base current would be still in the exponential increase region. Furthermore, it is noted that electric field caused by high emitter series resistance enhances the defect-assisted current components just explained above.

3.3 Improvement of Current Gain

It can be said from above considerations that the current gain degradation is occurred due to low current supply capability of the emitter rather than the interface problem.

In order to confirm this hypothesis, as an example, a micro-crystalline Si(μc-Si) emitter which provided wide band gap as same as a-Si [3] in spite of much lower resistance was used. The μc-Si was also prepared by L-couple plasma CVD apparatus, under the same condition as deposition of a-SiC film except adding Ar

gas of 20% to SiH_4, PH_4 gas sources. In the case without Ar gas, the film was a-Si phase, not being micro-crystallized. The conductivity, band gap, grain size and film thickness of μc-Si is 4.4 S/cm, 1.75 eV, about 50 Å and 100 nm respectively. Collector current density - current gain characteristics of the samples are shown in Fig.6. The current gain as high as 480 is obtained. This value is from 2 to 3 times larger than the value obtained by a-SiC emitter. Therefore we can say that this experimental result confirms the validity of the above considerations.

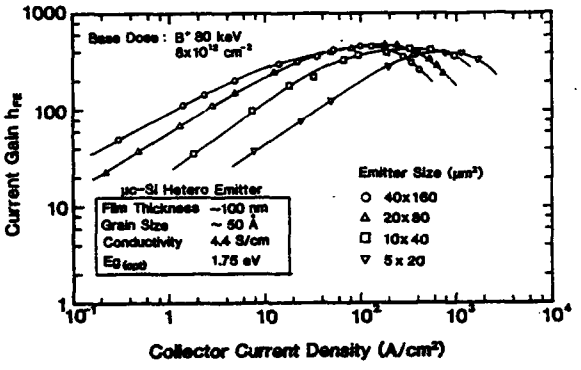


Fig.6. h_{FE}-I_C characteristics of μc-Si hetero-emitter transistors.

4. CONCLUSION

We have discussed the current transport mechanisms and current gain in a-SiC emitter Si-HBT. It can be concluded that the current gain of this type HBT is expected to be more improved by increasing the current supply capability of the emitter.

As a matter of course, the hetero-interface properties also should be improved. In order to realize this and obtain a high quality amorphous or micro-crystalline SiC film, we think that the films are necessary to be densified.

ACKNOWLEDGMENTS

The authors wish to thank Assoc. Prof. H. Ishiwara and Dr. Asano for useful discussions and Mr. Fukazawa for his technical support.

This work was partially supported by CASIO Science Promotion Foundation and 1987 Grant-in-Aid for Encouragement of Young Scientist from the Ministry of Education, Science and Culture.

REFERENCES

[1] K. Sasaki, S. Furukawa and M. M. Rahman : 1985 Int. Elec. Dev. Meeting, Washington D.C., Tech. Dig. pp. 294-297.

[2] K. Sasaki and S. Furukawa : 18th Int. Conf. Solid State Device and Materials, Toky, pp. 291-294 (1986).

[3] M. Ghannam, J. Nijs, R. Mertens and R. DeKeersmaecker : 1984 Int. Elec. Dev. Meeting, Sanfrancisco, Tech. Dig. pp. 746-748.

DEVICE DESIGN OF A HIGH VOLTAGE BiCMOS IC

Tong, Qin-Yi and Wu, Wei

Microelectronics Center
Nanjing Institute of Technology
Nanjing 210018, China

A new high voltage (H.V) npn bipolar transistor for H.V BiCMOS IC's has been developed which is fully compatible with conventional low valtage n-well CMOS process. The npn transistor employs n-well of low valtage (L.V) CMOS as the collector drift region and it acts as the self-isolation region as well. The narrow self-aligned base is a result of double diffusion. The device has shown a high performence, i.e, h_{FE} of 100, f_T of 31 MH_Z and BV_{ceo} of greater than 300 V. It is expected that complementary H.V npn and pnp transistors of this type can be integrated with H.V and L.V CMOS and bipolar devices on a same chip by Silicon wafer Direct Bonding (SDB)/ SOI technology.

1. INTRODUCTION

Many ASIC (Application Specific Integrated Circuit) designs require high voltage devices to be integrated compatibly in a same chip with conventional low voltage (L.V) control and drive circuits. Moreover, high voltage (H.V) IC's have been demanded not only by digital applications, such as flat display panels but also by analogue circuits such as interface circuits in telephone systems and piezoelectric actuator driver circuits. For H.V digital applications, high driving capability and low power are main requirements while H.V analogue circuits with high gain band width and low noise are most desirable. Because of the inherent advantages of CMOS to digital and bipolar to analogue performence, the development of a compatible process technology combining L.V CMOS and bipolar with H.V CMOS and bipolar devices on one chip provides an ideal solution to satisfy all the requirements. A monolithic integration technology of L.V CMOS/bipolar with H.V NMOS has been reported by us previously (1), (2). Based on the same n-well CMOS process, a new H.V biploar device is proposed which utilizes the n-well drift region and emitter extended field plate to provide high breakdown voltage BV_{ceo}. The same n-well provides self-isolation between devices and the narrow base is realized by double diffusion technique for high gain and high speed.

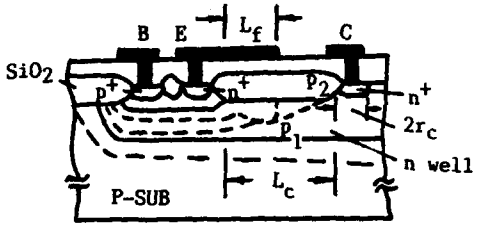
The device structure and model is presented. Experimental results show the proposed H.V bipolar device has high performance and is suitable for H.V BiCMOS IC's. Monolithic integration of L.V CMOS/complementary bipolar with H.V CMOS/complementary bipolar can be realized by newly developed Silicon wafer Direct Bonding technique, ie, (SDB)/SOI technology and it provides maximum design flexibility for various system applications(3),(4).

2. DEVICE STRUCTURE AND MODEL

Fiqure 1 is the cross-section of the new H.V npn device. The 4.5 μm deep n-well of L.V CMOS is adopted as a self-isolated collector and the source/drain diffusion forms n^+ collector contact and n^+ emitter which is self-aligned to the base edge. An additional photo step to open the base region, a boron implant and a drive-in step to form the base are performed. The aluminium or poly emitter contact acting as a field plate is across over the thick field oxide and the drift region to increase the device breakdown voltage BV_{ceo}.

As V_{ce} increases, the depletion layer on the n-well surface and that in the inside bottom of the n-well, caused by the field plate potential and the reverse biased voltage on the

well, respectively, spread downward and upward, respectively, to meet each other. At the voltage of V_{p1}, the well is pinched-off before the electric field in base-collector junction reaches a critical avalenche electric field. The further increased valtage is mainly dropped in the pinch-off region and causes another pinch-off to occur close to the n^+ end of the well, P_2 at voltage of V_{p2} until the breakdown of the well junction takes place.



Fiqure 1 Cross-section of H.V npn device structure

V_{p1} is given by

$$V_{p1}=\left\{-\left(\frac{qN_B}{2\varepsilon_o\varepsilon_{si}}\right)^{\frac{1}{2}}\left(x_{jw}+\frac{\varepsilon_{si}t_F}{\varepsilon_{sio_2}}\right)+\frac{1}{2}\left(\frac{2qN_B}{\varepsilon_o\varepsilon_{si}}\left(x_{jw}+\frac{\varepsilon_{si}t_F}{\varepsilon_{sio_2}}\right)^2+\frac{2qN_w}{\varepsilon_o\varepsilon_{si}}\left(x_{jw}^2+\frac{2\varepsilon_{si}t_F}{\varepsilon_{sio_2}}x_{jw}+\frac{2\varepsilon_o\varepsilon_{si}}{qN_w}V_{FB}\right)^{1/2}\right)^2\right\} \quad (1)$$

Where N_B and N_w are doping concentration of p-substrate and n-well, respectively, t_F thicknees of the thick oxide, x_j the junction depth of n-well, V_{FB} the flat band voltage of thick oxide capacitor.

V_{p2} is expressed by

$$V_{p2}=\frac{qN_w}{2\varepsilon_o\varepsilon_{si}N_B}(N_B+N_w)x_{jw}^2 \quad (2)$$

The collector on-resistance R_c is composed of drift region resistance beneath the field plate, R_1 and the resistance of the rest of the drift region, R_2:

$$R_1=\frac{\ln\left(1-\frac{L_f}{r_c+L_c}\right)}{2\pi q\mu_e N_w x_{jw}}\left\{\left(1+\frac{\varepsilon_{si}t_F}{x_{jw}\varepsilon_{sio_2}}\right)-\frac{1}{N_w}\left(\frac{2\varepsilon_o\varepsilon_{si}N_B}{q}\left(V_{p1}+\phi_{bi}\right)\right)+\frac{1}{x_{jw}}\left(\left(\frac{\varepsilon_{si}t_F}{\varepsilon_{sio_2}}\right)^2+\frac{2\varepsilon_o\varepsilon_{si}}{qN_w}\left(V_{p1}-V_{FB}\right)\right)^{\frac{1}{2}}\right\}^{-1} \quad (3)$$

$$R_2=\frac{\ln\left(1+\frac{L_c-L_f}{r_c}\right)}{2\pi q\mu_e N_w x_{jw}}\left\{1-\frac{1}{N_w}\left(\frac{2\varepsilon_o\varepsilon_{si}N_B}{q}\left(V_c+\phi_{bi}\right)\right)^{\frac{1}{2}}\right\}^{-1} \quad (4)$$

where round shape of the device is assumed and V_c is the collector voltage.

To optimize the design, a trade off between BV_{ceo} and R_c is made. Fiqure 2 is the experimental results of L_c vs. BV_{ce}. It is clear that for a specific required BV_{ce}, the shortest possible L_c should be chosen.

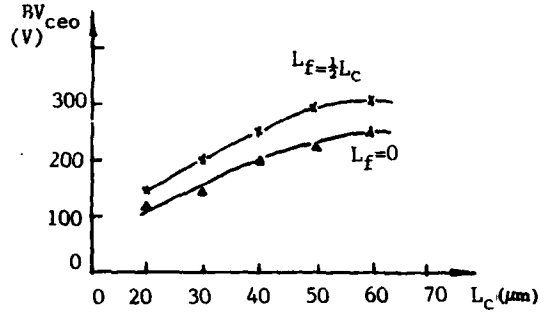


Figure 2 Experimental relationship between BV_{ceo} and L_c

3. EXPERIMENTAL RESULTS

The device has been successfully proecessed by our n-well CMOS process with one extra mask.

Fiqure 3 is I-V characteristres of the device. Because of the field plate stracture, BV_{ceo} is not a function of current gain h_{FE} and the base punch-through is less a problem.

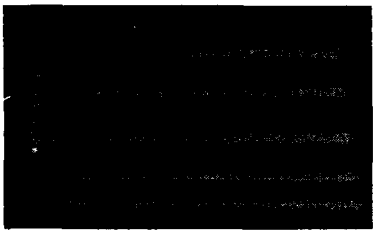


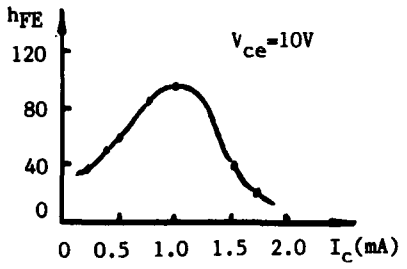
Figure 3 I-V characteristic of H.V npn transistor

Horizontal axis, V_{CE}, 50V/div;

Vertical axis, I_c, 500μA/div;

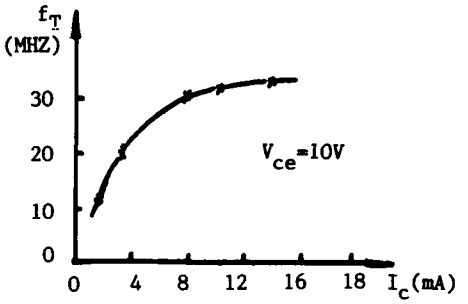
Base current, 5μA/step

Figure 4 shows the experimental result of h_{FE} vs. collector current I_c. The device has about 6500 μm^2 emitter area and L_c of 70 μm. It is expected the degradation of h_{FE} at higher current will be improved by optimizing the device design and process parameters.



Fiqure 4 Experimental relationship of h_{FE} vs. I_c

High f_T is obtained, as can be seen from Fignre 5, by small junction area with no degradation of breakdown performance. From Fiqure 3-5, it is clear that the fabricated H.V npn transistor has high performance, such as h_{FE} of 100 f_T of 31MH_Z and BV_{ceo} of grater than 300V, eventhough much improved results can be expected by furthur optimization.



Fiqure 5 Experimental results of f_T VS. I_c

4. CONCLUSION

A new H.V npn transistor for H.V BiCMOS IC has been developed. The main feature of this H.V BiCMOS technology is full compatibility with a conventional VLSI CMOS process. It leads to low cost, better performance and high flexibility of ASIC designs. As advances of new isolation technologies, for example, SOI substrate produced by SDB technique, complementary H.V npn and pnp transistors of this kind can be integrated with L.V CMOS/bipolar on a same chip.

We are grateful to staff of MCNIT for their process service and the work is supported by Chinese Natural Science Foundation(6866011)

REFERENCES

1. Tong, Q-Y,"CMOS Interfac Technology for VLSI Systems", International Vacation School on VLSI Fabrication, Edinburgh, U.K, April, (1984)

2. Tong, Q.-Y, etal,"Process and Device Design of A Fully Compatible 300-Volt CMOS IC", Jurnal of Nanjing Institute of Technology, Vol.2, (1985)

3.Lasky,J.B, etal,"Silicon-on-Insulator(SOI) By Bonding and Etch-Back", Tech. Dig. of IEDM, (1985), 684

4. Li Hui, etal,"Some Material Structure Properties of SOI Substrate Produced by SDB Technology", INFOS'87, Leuven, Belgium, April,(1987)

Session B4.3

Ultra Fast Optoelectronics

Chairman: H. Melchior

Thursday, September 17, 1987

LASER DIODE WITH AN INTEGRATED GAIN/LOSS MODULATOR FOR THE GENERATION OF PICOSECOND OPTICAL PULSES BY ACTIVE MODE LOCKING

J. Werner, G. Guekos and H. Melchior

Swiss Federal Institute of Technology
Institute of Quantum Electronics
CH-8093 Zurich, Switzerland

Lasers with an integrated modulation section have been used to generate ultra-short actively modelocked optical pulses with 8 ps duration and powers in excess of 400 mW. The pulses are of smooth shape and free of substructures and bursts.

1. INTRODUCTION

The generation of ultrashort light pulses by semiconductor lasers is of interest for various applications. Mode locking techniques - active, passive or a combination of both - have been applied to diode lasers in external cavities by several groups to produce repetitive pico- and even subpicosecond pulses /1/. In many cases, however, the analysis reveals bursts of individual pulses or, in the case of single pulses, pulse shapes that suffer from noise and other instabilities. Recently subpicosecond pulses have been obtained from passively mode locked diode lasers with external multiquantum well (MQW) absorber but with an additional external pulse compression stage /2/. The purpose of this work is the realization of a special laser diode and its implementation in a compact setup for the generation of ultrafast repetitive optical pulses. We report the development of an actively mode locked GaAs/AlGaAs Transverse-Junction-Stripe (TJS) laser with an integrated gain and loss modulation section. The laser was operated in an external cavity with a diffraction grating. To the best of our knowledge, this is the first time that such a configuration was used to generate actively mode locked repetitive pulses of

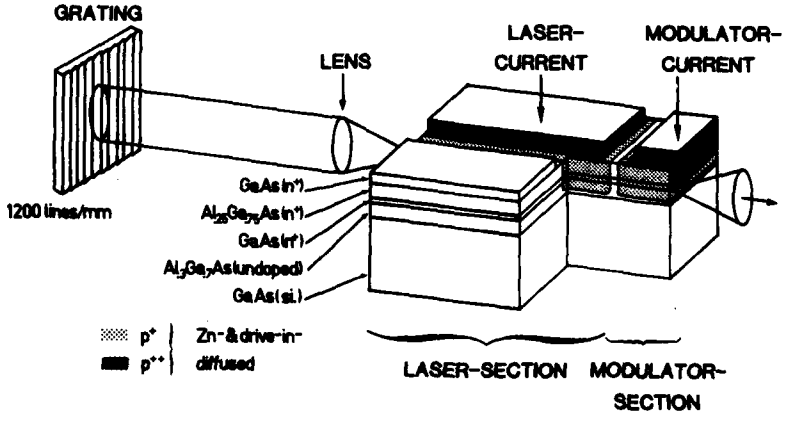


FIGURE 1

Partial cut-view of the laser diode in an external cavity setup with diffraction grating.

smooth shapes with durations as short as 8 ps and powers in excess of 400 mW. The pulses are free from bursts and substructures due to spurious reflections. An additional feature is the enlarged frequency tuning range within which the pulses remain short and free of distortion.

2. EXPERIMENTAL SETUP AND DEVICE FABRICATION

The mode-locking setup with the semiconductor laser, the modulator and the external cavity arrangement is shown schematically in fig. 1. Both the laser and the modulator are fully integrated transverse-junction AlGaAs heterostructures. These planar four layer heterostructures are grown by LPE on semiinsulating GaAs substrates. A masked double Zn-diffusion is used to form the lateral p^{++}-p^{+}-n^{+} doping profiles of the junction and index-guiding structures of the laser and the modulator. The first diffusion step was performed at 700°C for two hours using a specially developed homogeneously sintered and equilibrated Zn_3As_2 - $Zn As_2$ - GaAs source /3/. The second step consisted of a drive-in diffusion carried out at 850°C for 90 minutes in an evacuated, sealed quartz ampoule. During the second diffusion step the surface of the chip was protected by a 3000 Å thick SiO layer. After the chip is contacted from the top side the upermost GaAs layer is etched away in between the contacts, to prevent by-pass currents from flowing through the otherwise present parasitic p-n junction. With the separate diffusions the laser section of 210 µm length is electrically isolated from the 30 µm long modulator section. Both sections are optically interconnected by an integrated 20 µm long AlGaAs-waveguide whose coupling efficiency is estimated to be better than 95%. The wafer is thinned down to 40 µm and the device is mounted junction-up with Indium solder onto a type IIa diamond heat sink. To minimize the reflection losses the laser facet facing the external cavity is antireflection coated with a $\lambda/4$ SiO layer. The residual reflectivity of this facet is estimated from the visibility fringes of the spectrum to be lower than 10^{-3}.

3. EXPERIMENTS

For mode locking the laser is placed into a grating-terminated external cavity with a length of 30 cm. Light emerging from the laser facet is collimated by an antireflection coated N.A. $\sim$ 0.6 lens and directed onto a diffraction grating with 1200 lines/mm that was blazed at 1 µm. The beam diameter perpendicular to the lines of the grating was 1 mm. From the grating, which is tilted by 32.6° with respect to the beam direction, the light is reflected and coupled back into the laser. The overall coupling efficiency is estimated to be 10% or higher. In operation the laser, whose threshold is about 65 mA is driven by a dc current in the range of 70-85 mA. The modulator section is forward driven from a comb generator that is connected via a 50 Ohm line. The electrical pulses flowing through the modulator section have a peak value of 215 mA and a FWHM of 130 ps. The light output emerging from the modulator facet is used for pulse measurements, autocorrelation and spectral analysis.

The autocorrelation trace of the periodic pulses obtained by second harmonic generation (SHG) in a $LiJO_3$ crystal are shown in fig. 2.

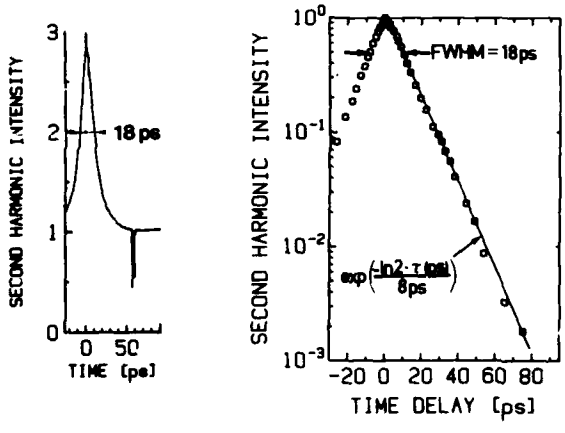


FIGURE 2a) Linear plot and 2b) semi-log plot of the autocorrelation function (ACF) of the pulse intensity. Note the exponential dependence over almost three decades. The symmetrical shape of the ACF is confirmed experimentally whereas the figure shows in detail only the part for positive time delays.

For these measurements, the laser was prebiased to a dc current of 80 mA. The comb generator is adjusted to 494.8 MHz, the frequency for optimally short optical pulses. A somewhat higher maximum of the SHG output is observed for 496.216 MHz. As can be seen from fig. 2a the SHG traces are smooth and show the 3:1 peak to background ratio expected for background free pulses. A semilog plot (fig. 2b) reveals a smooth and precise exponential dependence on the delay time over almost three decades of SHG-intensity. This indicates that the pulses are extremely stable and free from disturbances due to spurious reflections within the external cavity. Assuming the individual pulses to be of one-sided exponential shape /1,4/ on determines from the FWHM of the SHG and the exponential shape of the SHG the width of the actual pulses to be as short as 8 ps. The peak power of these pulses was 432 mW. Fig. 3 shows how the FWHM of the autocorrelation function (ACF) and pulse energy depend on

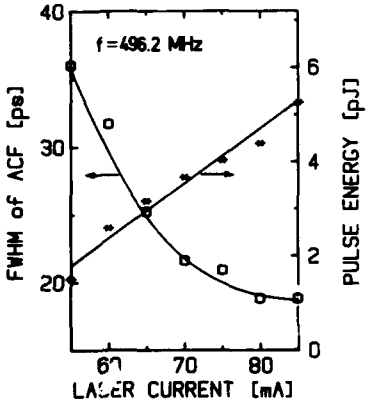


FIGURE 3

FWHM of the autocorrelation function (ACF) and pulse energy v.s. laser dc current.

the dc current through the gain section. At lower laser currents the pulse energy decreases and the pulses become wider. Actually the shape of the ACF is not as perfectely exponential and smooth as shown in fig 2b. The laser emission is centered at 8984 Å (see fig. 4). The width of the spectrum measured under the operating conditions of fig. 2 is 0.3 ± 0.04 nm which corresponds to 110 ± 15 GHz. The resulting time bandwidth product of 0.9 is about eight times larger

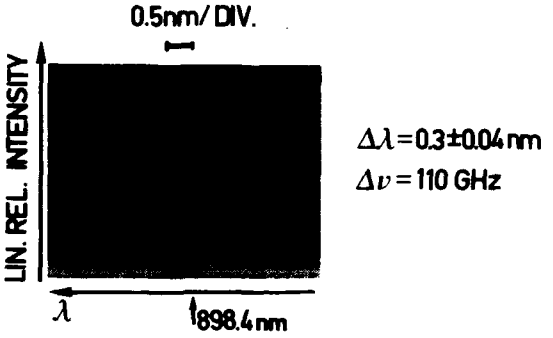


FIGURE 4

Optical spectrum of the pulse of figure 2a).

than the predictions for transform limited pulses /1/.

The repetition rate of the pulses is in essence determined by the length of the external cavity. It was about 496 MHz. By means of the applied signal the repetition frequency can be tuned externally. The tunability without adverse effects on pulse shape and duration was about ± 1 MHz. This twin section laser structure thus allows relative tuning ranges that are one to two orders of magnitude larger than usually encountered in ps laser setups /5/.

4. CONCLUSIONS

We have demonstrated the suitability of TJS-GaAs lasers with electrically isolated gain and modulator sections for the generation of stable picosecond pulses without substructures and bursts. The lasers are actively mode locked in an external cavity and deliver 8 ps pulses of over 400 mW at 898 nm and 496 MHz. In practical applications the large detunability eases the requirements on frequency precision of the modulating generator.

REFERENCES

/1/ Van der Ziel, J.P., Semicond. and Semimet., 22, part B, (1985) pp. 1-68.

/2/ Silberberg, Y., Smith, P.W., Subpicosecond Pulses from a Mode-Locked Semiconductor Laser, IEEE J.of Quantum Electron., Vol. QE-22 (june 1986) pp. 759-761.

/3/ Werner, J. et al, Zinc-Diffusion in GaAs Using a Sintered Ternary Source, Jap. J. Appl. Phys. Vol. 26, No. 4, (April 1987) pp. 641-642.

/4/ Bessonov, Yu. et al, Sov.J.Quantum Electron., 15, (1985) pp. 435-436.

/5/ Bradley, D.J. et al, IEEE J.of Quantum Electron., QE-17, (1981) pp. 658-670.

INDIUM TIN OXIDE - GALLIUM ARSENIDE PHOTODIODES FOR OPERATION AT FREQUENCIES BEYOND 110GHz

D G Parker and W Sibbett*

GEC Research Limited, Hirst Research Centre, East Lane, Wembley, Middlesex, HA9 7PP, United Kingdom

* Dept of Physics, University of St Andrews, St Andrews, Fife, KY16 9SS, United Kingdom

An Indium Tin Oxide/GaAs photodiode has been demonstrated with -3dB bandwidths in excess of 110 GHz (~4 ps FWHM). This device exhibits an external quantum efficiency of >25% (~0.2A/W) at 820 nm and is fundamentally limited by the active layer thickness. The device has been mounted such that direct measurements of bandwidth could be made using external mixers and a high-frequency spectrum analyser. The bandwidth agrees well with the Fourier transform of a theoretically predicted pulse duration of 4.2 ps FWHM. Corroborative data obtained using the optoelectronic cross-correlation technique are also presented.

1. INTRODUCTION

A fundamental component in any high frequency optical system is a low noise photodetector with sufficient bandwidth to meet the system requirement. In recent years, with the advent of short pulse lasers, directly modulated high frequency semiconductor lasers, external modulators and laser heterodyne techniques, there exists a need for fast photodetectors so that powerful system concepts can be realised.

Devices such as radiation damaged photoconductive detectors have been demonstrated with response times in the order of 4 ps [1]. However, these devices are extremely inefficient normally requiring very high optical powers to achieve usable electrical signals. A different approach is the use of a vertical geometry detector such as a p-i-n or Schottky barrier photodiode: the latter having the added advantage of a reduced minority carrier contribution associated with the p-layer. These structures enable the absorption layer thickness to be engineered to obtain the optimum compromise between capacitive and carrier transit time effects.

It has been demonstrated [2] that, by using a thin, semi-transparent metal to form a Schottky barrier photodiode, bandwidths in the order of 100 GHz can be achieved. These devices unfortunately suffer from relatively low photosensitivities due to the limited transmission property of the metallic layer. A solution is to use conducting Indium Tin Oxide (ITO) to form the rectifying contact to GaAs. It has been shown previously [3] that this material produces diodes of excellent quality capable of detecting light at very high frequencies whilst transmitting the majority of the incident radiation.

In this paper we report the fabrication and assessment of devices operating beyond 110 GHz whilst exhibiting a high quantum efficiency of 25% for 820 nm radiation. We believe this not only to be the fastest photodiode, but that it also exhibits the highest responsivity -bandwidth product yet reported. In addition, this is the first time a detection bandwidth of this magnitude has been measured directly and the results compared with optoelectronic correlation data.

2. DEVICE FABRICATION AND D.C. CHARACTERISATION

A schematic representation of the device structure is shown in Figure 1. It consists of a 0.45 μm thick n^-absorbing layer ($N_D = 5 \times 10^{15}cm^{-3}$) over an underlying 0.5 μm n^+ ($N_D = 3 \times 10^{18}cm^{-3}$) buffer layer to aid in

the formation of a low resistance ohmic contact. The GaAs material is grown via the Vapour Phase Epitaxy technique at 715°C giving a growth rate of 0.2 μm min^{-1}. Delination of the device structure is achieved by a combination of wet mesa etching and proton damage. Hence, as the device structure is grown on semi-insulating substrate the parasitic bonding pad capacitance becomes negligible. The active layer is 5 μm x 5 μm, resulting in a capacitance of <10 pF.

The ITO layer thickness is chosen to exploit the dielectric properties of the material. The refractive index of the deposited layers is ~2 making it an ideal antireflection coating to GaAs. Hence, a 100 nm layer is deposited of composition $In_{0.9}Sn_{0.1}O$ by R.F. sputtering from an ITO target.

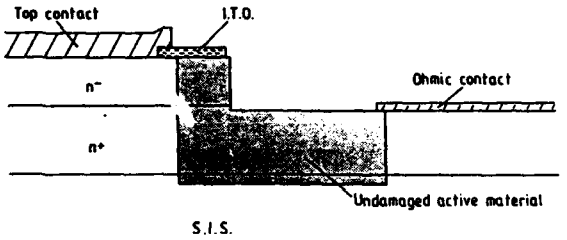


Figure 1: Schematic section through the device

A D.C. electrical characterisation of the structures was carried out using both current/voltage and capacitance/voltage techniques. The barrier heights and ideality factors of typical ITO/GaAs diodes are (0.85 ± 0.05)eV and (1.05 ± 0.05) respectively. ITO is a wide bandgap semiconductor with a bandgap Eg greater than 3.75 eV with its conductivity arising from n-type behaviour. In terms of electrical properties, therefore, the material may be regarded as being degenerately doped whereby it exhibits semimetallic characteristics and a Schottky barrier is thus formed. The reverse leakage currents are typically <5 nA and sub-nanoamp values have been measured for a bias voltage of -5V. At their normal operating voltage of -3V the devices have a measured quantum efficiency as high as 25% (0.2 A/W) at 820 nm which is limited by the active layer thickness.

3. DEVICE AND PACKAGE DESIGN

A theoretical design model has been developed which simulates both the inherent device response and the perturbation imposed by the external circuit. The carrier dynamics are modelled by solving Poisson's equation for the charge density inside the device with appropriate injection and boundary conditions associated with the applied bias and the inherent barrier height. This then gives the localised electric field which will cause carrier flow and hence current. The new charge distribution is then calculated using the continuity equation. Having obtained this response the SPICE equivalent circuit routine is used to model the effects of capacitance and inductance. A simulation for the 5 μm x 5 μm square device is shown in Figure 2. The series package inductance used was 100 pH. The FWHM of this pulse is 4.2 ps.

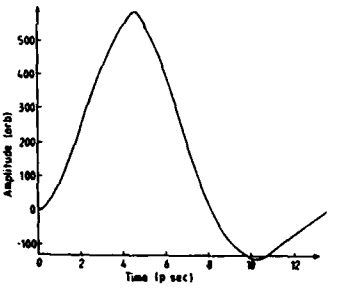


Figure 2: Theoretically predicted impulse response showing FWHM of 4.2 ps

This 100 pH inductance is incorporated into the package design which is critical. The device is mounted on a quartz substrate, microstrip submount. The bias is applied to the device via a hybrid integrated bias network. This consisted of a high R.F. impedance line of $n\lambda/4$ filtering length. The D.C. return path for the bias is via the waveguide casing.

The signal is transferred into an appropriate gauge waveguide for the frequency

band of interest by a suitably designed millimetric transition. Such a transition routinely demonstrates an insertion loss of 0.4 dB and a return loss of > 20 dB over the 75-110 GHz frequency band when referred to a callibrated signal source.

4. HIGH FREQUENCY OPTICAL ASSESSMENT

Once mounted in the package described above the bandwidth of the detector can be assessed. The output is directly fed into a pre-calibrated, waveguide mounted mixer diode which uses the amplified local oscillator signal from a HP spectrum analyser as a reference signal for the mixer. The analyser is then able to display the amplitude of the detector output against frequency.

The optical source used was an argon-pumped, colliding pulse mode-locked dye laser. Incorporated into this passively mode-locked C.W. ring configuration is an intracavity sequence of four Brewster-angled prisms which compensate for group velocity dispersion effects allowing stable femtosecond pulses stable femtosecond pulses to be established. The laser routinely demonstrates a pulse duration of < 50 fsec at a repetition rate of 100 MHz and at an operating wavelength of 630 nm.

The Fourier Transform of such a chain of pulses has an amplitude envelope which is the transform of one individual pulse. Within this exists a comb of real components which are spaced by the cavity frequency. In this case this will be -3dB flat to over 1 THz and hence any observed roll off will be associated purely with the detector under test after the data is normalised to the microwave test system. Hence, this measurement technique is absolute in the result obtained and it is simple to extract the data. This is not the case in some correlation measurements, such as that described below, where the calculated response depends upon the assumed pulse shape or in the more elaborate electro-optic sampling methods involving a Pockel's cell for example.

The measured response characteristic in the 75-110 GHz band for an applied reverse bias of

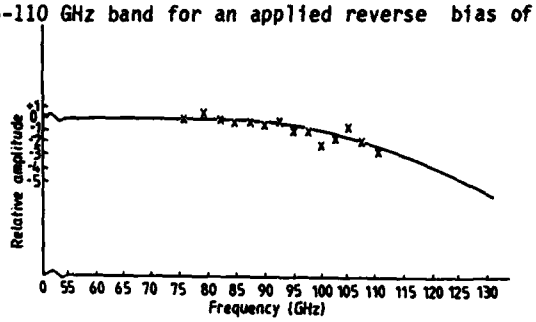


Figure 3: Frequency response of detector showing experimental data (crosses) and theoretically predicted characteristic (solid line).[The theoretical fit is the FFT of Figure 2].

-3V is shown in Figure 3 along with a theoretical fit which is the Fast Fourier Transform of the impulse shown in Figure 2. The response is flat in the region 0-75 GHz within ±1 dB. It is clear from this data that there is excellent agreement between theory and experiment, and that the -3 dB bandwidth is around 115 GHz which is extrapolated from the available data. This close agreement would imply that the FWHM impulse is approximately 4 ps.

5. CORRELATION MEASUREMENTS

The use of an optoelectronic correlation system in the assessment of high speed detectors was first demonstrated by Auston and Smith [4]. Either two fast photoconductive sampling gates are used to obtain an autocorrelation or one gate is used to sample the response from a photodiode in a cross-correlation. To determine the response time of the ITO/GaAs photodiode a correlation circuit response was constructed from 50 Ω track on a quartz substrate to limit dispersion. In this experiment higher powers were required and hence, a synchronously pumped dye laser was used with a frequency doubled C.W. mode-locked Nd:Yag laser as the pump

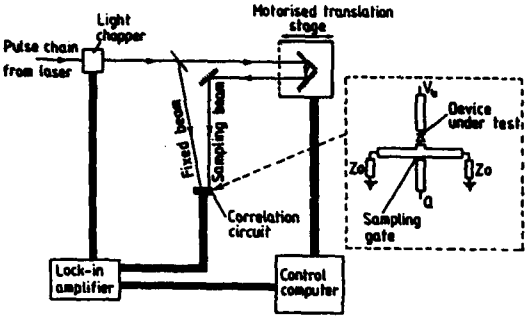


Figure 4: Experimental arrangement for correlation measurements

source. This laser produced 3 ps pulses with an average power of 100 mW at 620 nm at a repetition rate of 82 MHz. The overall experimental arrangement is shown in Figure 4.

Initially, two proton damaged GaAs photoconductors were used. The damage dose was 10^{14} at an energy of 180 keV and is introduced into the sample to reduce the free carrier recombination time. The correlation function is shown in Figure 5(a). Reversal of the sample and signal beams yields a circuit element response in the order of 1 ps. If, Gaussian pulses are assumed, analysis of the response results in a sampling aperture of 4.5 ps. This value is suspect as the functions are not purely Gaussian. The measurement of the device under test actually perturbs the system, as in an electrical correlation the two pulses are not isolated. Hence, the function is not a true auto-correlation and as a result the Gaussian approximation is made in the absence of the real pulse functions.

Figure 5(b) shows the correlation between the switch and a 5 μm x 5 μm ITO photodiode, a similar analysis yielding a FWHM response time of 4.3 ± 2 ps for the photodiode where the error in this result is an estimated order of the data reliability from the approximation used. This value is very close to that previously discussed and acts as a confirmation of the response. It is felt, however, that the directly measured bandwidth is more reliable

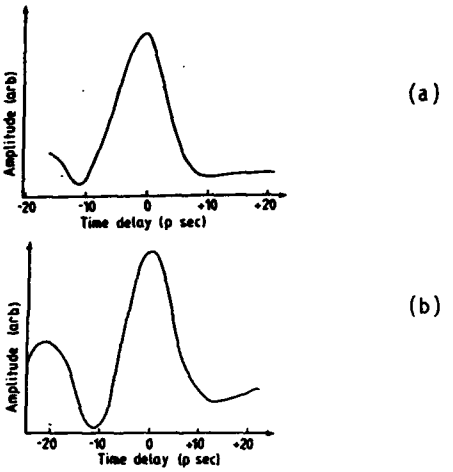


Figure 5: Correlation function for (a) two GaAs gates and (b) a 5 μm x 5 μm ITO/GaAs photodiode with a GaAs sampling gate.

and that the extremely close agreement is perhaps fortuitous.

6. CONCLUSIONS

A 5 μm x 5 μm photodiode has been designed, fabricated and assessed. The device exhibits a high quantum efficiency of 25% (0.2 A/W) at 820 nm due to the use of transparent layer of Indium Tin Oxide to form the rectifying contact. The -3 dB bandwidth has been directly measured to be greater than 110 GHz (~4 psec FWHM) which is in good agreement with a theoretical prediction and conventional correlation data.

7. ACKNOWLEDGEMENTS

The authors acknowledge the technical assistance of P Say, A Hansom, K Enfield, A Finch, G Chen, L Teale and J Sabir

8. REFERENCES

1 Hammond R.B., Paulter N.G. and Wagner R.S., Appl. Phys. Lett. 45, p289, (1984)
2 Wang S.Y. and Bloom D.M., Elect. Lett., 19 p554, (1983)
3 Parker D.G., Elect. Lett., 21, p778, (1985)
4 Auston D.H. and Smith P.R., Appl. Phys. Lett 41, p599, (1982)

HIGH-SPEED CONTROL OF MICROWAVE SIGNALS USING InP:Fe PHOTOCONDUCTIVE DEVICES

Ingmar ANDERSSON and Sverre T. ENG*

Department of Electrical Measurements
Chalmers University of Technology
S-412 96 Gothenburg, Sweden

The feasibility of using InP:Fe photoconductive devices as high-speed microwave switches has been demonstrated in the 0.01 - 10 GHz frequency range, by fabricating a fiber optic compatible modified interdigitated gap (MIG) structure. Rise and fall times of less than 100 ps have been measured and a power switching ratio (PSWR) of 13 dB at 10 GHz has been achieved when illuminating the device with 10 mW of effective optical power (CW) from a semiconductor laser with a fiber pigtail.

1. INTRODUCTION

Photoconductive devices have gained much attention for high-speed switching of d.c. or r.f. signals up to the gigahertz range [1-6]. This is due to their picosecond response times, high power handling capability, simplicity of operation and inherently near-perfect isolation of electrical and optical signals. The introduction of these optoelectronic devices in microwave systems may lead to new applications where high-speed, low weight, and low cost are of outmost importance e.g. future airborne radars.

Since the speed is mainly limited by the lifetime of the photo-induced carriers, the choice of semiconductor material is crucial. Methods for short pulse generation in materials with long lifetimes (e.g. i-Si) have been demonstrated [1-3], but these techniques suffer from limited repetition rate and the need for two mutually delayed optical pulses. Devices made from materials with short carrier lifetimes such as GaAs:Cr and InP:Fe, do not need two optical pulses since they turn off automatically because of picosecond lifetimes [4-5]. Both these materials are capable of high-speed operation. However, since higher conductance values have been obtained in InP:Fe [5], this material is more promising for high-speed optoelectronic switching of microwave signals.

We report on an investigation of InP:Fe photoconductive devices as high-speed switches in the 0.01 - 10 GHz frequency range. A model for analyzing the microwave performance is presented. In contrast to other reported optoelectronic microwave switches [2-3], we have been using a novel device structure. The structure, which we call the modified interdigitated gap (MIG) structure, is compatible with fiber optic illumination and will improve the device characteristics at high frequencies.

2. MICROWAVE ANALYSIS

Figure 1a is a schematic presentation of the frequently used single gap (SG) microstrip structure and Fig. 1b is a top view illustration of the modified interdigitated gap (MIG) microstrip structure which we have used. The microwave performance can be analyzed by using the lumped element π equivalent circuit shown in Fig. 2, as a 1st-order approximation for frequencies up to 10 GHz [2,6]. The active region is modelled by gap and shunt conductances/capacitances (G_g, G_s, C_g, and C_s), where the conductances depend on both the optical intensity and wavelength. Furthermore, the narrow microstrip lines and the bonding wires (if used) are represented by inductances (L and L_B, respectively), and the outer gap by

* S.T. Eng has a joint appointment with Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA.

capacitances (C_{go} and C_{so}). In the case of SG devices the model reduces to that of the active region representation.

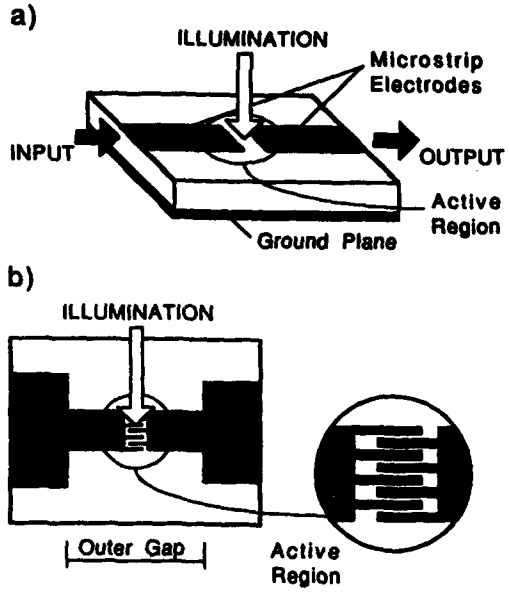


FIGURE 1

a) Schematic presentation of a single gap (SG) microstrip structure.
b) Top view illustration of the modified interdigitated gap (MIG) microstrip structure.

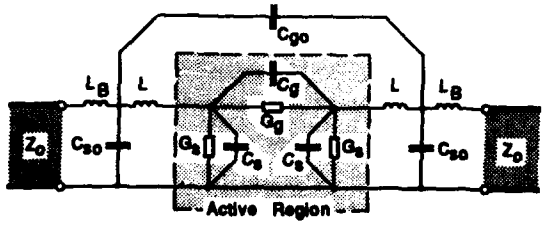


FIGURE 2

Equivalent π circuit model for analysis.

Referring to the model, the microwave transmission can be calculated from the transmission S-parameter S_{21}. The analytic expression of S_{21} in the general case (MIG) is quite complex. However, the basic characteristics can be seen from the expression for a SG device. The microwave power transmission coefficient for a SG device is given by (matched conditions, $L_B \approx 0$)

$$|S_{21}|^2 = \left| \frac{2Z_o(G_g+j\omega C_g)}{[1+Z_o(G_s+j\omega C_s)][1+Z_o(2G_g+G_s+j\omega(2C_g+C_s))]} \right|^2 \quad (1)$$

where Z_o is the characteristic impedance.

It is clear, from this expression, that the shunt elements will degrade the transmission characteristics. By using proper laser wavelengths, yielding a very thin photoexcited surface layer, the shunt conductances will be negligible small [6]. Moreover, the shunt capacitances can be minimized by keeping the gap length small compared with the thickness of the substrate [7]. Thus, the shunt elements can generally be neglected.

One of the most important parameters when using these devices as microwave switches is the power switching ratio (on/off), which will be (SG device)

$$PSWR = \frac{|S_{21}|^2_{on}}{|S_{21}|^2_{off}} = \left| \frac{2Z_o(G_{on}+j\omega C_g)}{1+2Z_o(G_{on}+j\omega C_g)} \right|^2 / \left| \frac{2Z_o(G_{off}+j\omega C_g)}{1+2Z_o(G_{off}+j\omega C_g)} \right|^2 \quad (2)$$

where G_{on} and G_{off} are the gap conductance in the illuminated and non-illuminated state, respectively.

Figure 3 shows the calculated power switching ratio versus frequency using some typical values for a SG device. It is clearly shown, that the main limiting factor at high frequencies is the gap capacitance. Consequently, in order to improve the PSWR, the gap capacitance must be reduced. This can be accomplished by using the MIG structure, as will be shown in the experimental results and discussions section. At low frequencies the PSWR is set by the ratio of G_{on} and G_{off}. High-resistivity materials such as InP:Fe will ensure low values of G_{off}.

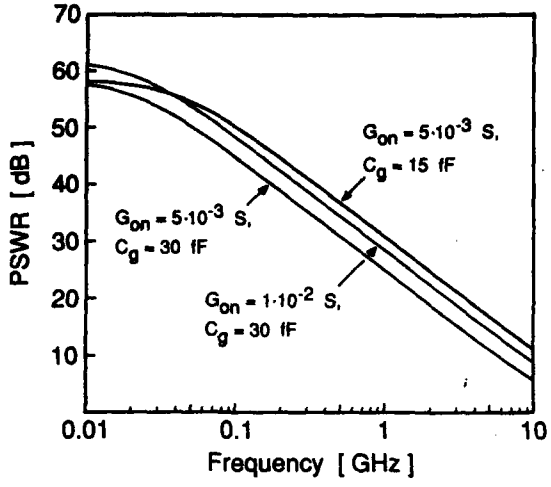


FIGURE 3

Calculated power switching ratio (PSWR) versus frequency for a SG device (Z_o = 50 Ω, G_{off}).

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

Photoconductive switches were fabricated on 400 μm thick slices of InP:Fe by conventional vacuum evaporation of contact material (Sn/Au) and lift-off techniques. For comparison purpose, both single gap (SG) and modified interdigitated gap (MIG) devices were made. The width of the microstrip lines (broad section on the MIG) was designed to maintain the 50 Ω geometry (300 μm). The electrode spacing (gap) was 4 μm, the interdigital electrode lines were 4 μm wide and 44 μm long (MIG). The overall active region for the MIG devices was 60 x 44 μm^2, and the outer gap was 200 μm. The slices were cut into individual devices (1 x 1 mm^2) and mounted (bonded) in a high-speed microstrip line test package.

The rise and fall times were measured by operating the devices as photoconductive detectors [1,4,5]. A directly modulated GaAlAs semiconductor laser producing optical pulses 90 ps wide (FWHM) was used for illumination. The electrical output shown in Fig.4, was monitored by a Tektronix sampling oscilloscope (t_r<25 ps). As can be seen in Fig. 4, the width of the electrical output is close to 90 ps (FWHM), indicating rise and fall times of less than 100 ps. The ringing is mainly caused by the microstrip test mount.

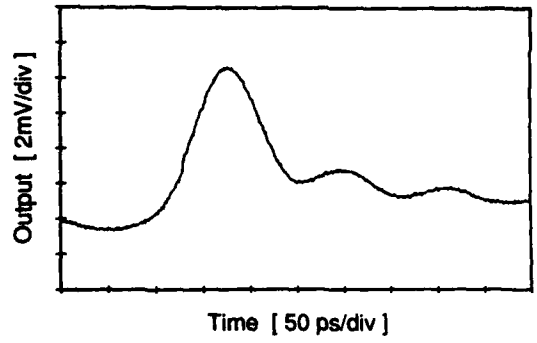


FIGURE 4

Measured pulse response from an InP:Fe switch.

The microwave performance was investigated in the 0.01 - 10 GHz range by measuring the power switching ratio (PSWR) using a tracking generator and a spectrum analyzer. A high-power GaAlAs laser, with a fiber pigtail (100 μm core) and CW-operated at 50 mW, was used for illumination. The fiber output was imaged on the active regions in a 1:1 scale by lenses. Figure 5 shows the measured power switching ratio for a MIG and a SG device. The corresponding theoretical curves have been calculated from the model in Fig. 2, using the data in Table 1. The good agreement between measured and calculated data implies that the model describes the performance well. The PSWR at 10 GHz for the MIG device was measured to be 13 dB, while no PSWR could be observed for the SG device, which clearly shows the advantages of the MIG structure. The difference in the PSWR at low frequencies is attributed to the different sizes in active regions (areas). The illumination efficiency (A_{active}/A_{fiber}) is about 20% for the MIG device and about 5% for the SG device, corresponding to an effective optical power of 10 mW and 2,5 mW, respectively.

TABLE 1. Device data

Device	G_{on} [S]	G_{off} [μS]	C_g [f]	C_s [fF]	C_{go} [fF]	C_{so} [fF]	L [nH]	L_B [nH]
MIG	5	10	3	.1	5	.9	.2	.5
SG	1.3	10	30	.01	-	-	-	.5

G_{on} and G_{off} are estimated from dc measurements.

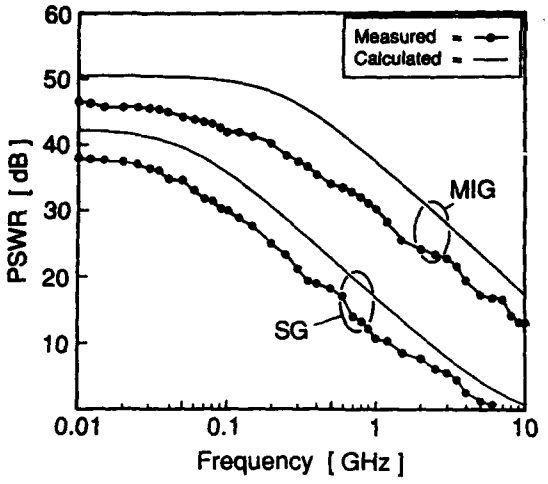


FIGURE 5

Measured and calculated power switching ratio (PSWR) versus frequency for a MIG and a SG device (Z_o = 50 Ω).

These results indicate that the modified interdigitated gap structure can be used as a high-speed microwave switch, for frequencies up to at least 10 GHz. By optimizing the device structure it would be possible to improve the power switching ratio at least by an order of magnitude and possibly extend the frequency range to 20 GHz.

4. CONCLUSIONS

The feasibility of using InP:Fe photoconductive devices as high-speed microwave switches has been demonstrated in the 0.01 - 10 GHz frequency range. The results show that the fiber optic compatible modified interdigitated gap (MIG) structure is advantageous when compared with the common single gap (SG) structure. Experimental data for the initially fabricated MIG devices indicate rise and fall times of less than 100 ps (limited by the optical pulse). Power switching ratios (PSWR) as high as 13 dB at 10 GHz has been measured when CW-illuminating the device with 10 mW of effective optical power from a laser with a fiber pigtail.

ACKNOWLEDGMENT

This work was supported by the National Swedish Board for Technical Development (STU).

REFERENCES

[1] D.H. Auston, "Picosecond optoelectronic switching and gating in silicon," *Appl. Phys. Lett.*, 26, pp. 101-103, 1975.

[2] A.M. Johnson and D.H. Auston, "Microwave switching by picosecond photoconductivity," *IEEE J. Quantum Electronics*, 11, pp. 283-287, 1975.

[3] W. Platte, "High-speed optoelectronic switching in silicon gap-shunt microstrip structures," *Electronics Letters*, 12, pp. 437-438, 1976.

[4] C.H. Lee, "Picosecond optoelectronic switching in GaAs," *Appl.Phys.Lett.*, 30, pp. 84-86, 1977.

[5] A.G. Foyt, F.J. Leonberger, and R.C. Williamson, "Picosecond InP optoelectronic switches" *Appl.Phys.Lett.*, 40, pp. 447-449, 1982.

[6] W. Platte, "Spectral dependence of microwave power transmission in laser controlled solid-state microstrip switches," *Solid-State and Electron. Dev.*, 2, pp. 97-103, 1978.

[7] M. Maeda, "An analysis of gap in microstrip transmission lines," *IEEE Trans. on Microwave Theory and Tech.*, 20, pp. 390-396, 1972.

EPITAXIAL SILICON AVALANCHE PHOTODIODES FOR SINGLE PHOTON DETECTION WITH PICOSECOND RESOLUTION

Sergio COVA, Giancarlo RIPAMONTI, Andrea LACAITA and Massimo GHIONI

Politecnico di Milano, Dipartimento di Elettronica, and
Centro Elettronica Quantistica e Strumentazione Elettronica CNR
P. Leonardo da Vinci 32, Milano 20133, Italy

Silicon avalanche photodiodes, designed to operate at bias higher than the breakdown voltage, have been fabricated in a p epitaxial layer over a p++ substrate. The device structure was designed to obtain ultrafast detection of single optical photons, as required in applications such as optical fiber testing, laser ranging, etc.. Experiments demonstrate a time resolution having 45 picoseconds full-width at half maximum and carrier diffusion effects remarkably lower than previous non-epitaxial devices.

1. INTRODUCTION

Ultrafast detection of single optical photons can be obtained by means of p-n junction devices operating in the triggered avalanche mode [1-9]. Uniform breakdown characteristics over all the active area are obtained with a suitable device geometry and a careful fabrication process. The junction can be biased above the breakdown voltage and a single photogenerated carrier can trigger a self-sustaining avalanche current, terminated by a quenching arrangement. By using active-quenching circuits [7,8] the devices are operated in accurately controlled conditions and with short deadtime. Remarkable performance is thus achieved in working conditions suitable for various applications, such as optical fiber testing, laser ranging, fluorescent decay measurements, etc. [9,10]. These devices, called single-photon avalanche diodes SPADs, are a solid-state alternative to photomultiplier tubes

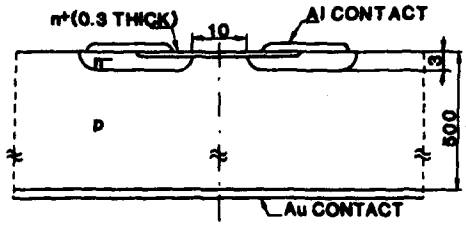


FIGURE 1

Schematic cross section of the previous SPAD devices (dimensions in microns).

providing wider spectral sensitivity range, higher resolution in the measurement of the photon arrival time, low detector noise. They are therefore particularly suitable detectors for the time-correlated single-photon counting technique [11]. Physical phenomena in the detector set limits to the time resolution and interesting problems of semiconductor device research have to be dealt with in order to improve it.

2. CARRIER DIFFUSION AND TIME RESOLUTION

The structure of the previously implemented SPAD devices [7-9] is sketched in Figure 1. A deep diffused n^- guard ring surrounds the shallow (0.3 μm) n^+ layer of the active junction. The ring diffusion avoids edge breakdown and provides a gettering action for the active junction region. The breakdown voltage V_b is about 28 V for devices fabricated in 0.6 Ohm cm p wafers.

The time resolution curve of a single-photon detector is given by the statistical distribution of the delay from the photon arrival time to the detection time [11]. It can be experimentally measured in a time-correlated single-photon counting set-up by using a source of ultrafast optical pulses. In our tests, gain-switched diode lasers [12] were used to generate pulses at wavelengths ranging from 785 to 904nm. The resolution curve of the device in Figure 2 exhibits two components: a fast peak and a slow tail, having intensity and shape dependent on

the photon wavelength. An example of experimental test is shown in Figure 2. The peak is due to photons absorbed in the depletion layer; full widths at half maximum (FWHM) values down to 60 ps have been reported by two of us [7-9]. The tail is due to carriers that are photogenerated in the neutral region beneath the junction and reach the depletion layer by diffusion. The wavelength dependence of the tail, which follows from the variation of the optical absorption coefficient, introduces a further complication in high resolution measurements. In fact, if the resolution function is accurately known, the shape of ultrafast optical signals can be obtained by deconvolution from the experimental data. However, in order to know the actual resolution function, it is necessary to measure, with good accuracy, the spectrum of the detected light and the resolution curves at the various wavelengths involved. This is a very complicate procedure, in practice not affordable in many cases. The design of new devices with improved performance appears to be a more interesting and effective approach, since the diffusion effects are strongly dependent on the device geometry [10].

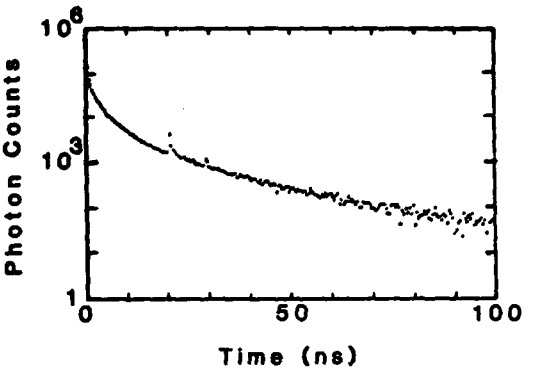


FIGURE 2

Resolution test on the previous SPAD device with laser pulses at 830 nm. The long time scale and the log vertical scale set in evidence the diffusion tail and compress the fast peak to a single point in the time origin with 4.1 10^5 counts.

3. SPAD DEVICE IN P EPILAYER OVER P^{++} SUBSTRATE

New SPAD devices were designed with the principal aim of improving the time resolution, that is, of reducing the weight of the diffusion tails and the width of the fast peak. A simple approach is to design a device structure in a p epitaxial layer over a p^{++} substrate, as shown in Figure 3. The basic idea is to have a) long carrier lifetime in the active region, as required for having a low thermal generation rate of carriers, and b) short lifetime in the substrate, in order to reduce the probability that photogenerated carriers diffuse out of it. A quantitative analysis of the time-dependent diffusion effects in the epitaxial device geometry was carried out by means of a Montecarlo simulation program, previously developed by two of us [10]. The computations showed that for an epistrate thickness of 12 μm and a junction geometry similar to the previous devices, a remarkable reduction of diffusion effects had to be expected provided that the substrate lifetime is sufficiently short, less than 100 ns. A 15 milliOhmcm substrate [12] was hence adopted. Experiments on the previous devices had shown that the width of the fast peak is reduced as the value of the maximum electric field is increased [9], denoting that the fluctuations in the avalanche build-up time decrease as the electric field is increased. The new devices were therefore designed to attain higher values of the maximum field, up to 550 kV/cm. A 4 Ohmcm epistrate was adopted; a boron ion implantation was used to increase the doping level in the

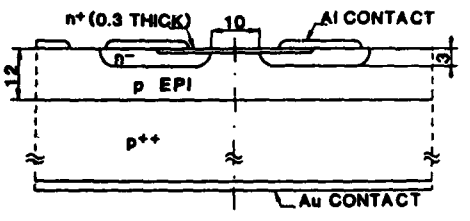


FIGURE 3

Schematic cross section of the SPAD device structure fabricated in a p epitaxial layer over a p^{++} substrate (dimensions in microns).

active junction, correspondingly increasing the electric field at breakdown and lowering the breakdown voltage to about 15 V. The guard ring was designed to have much higher breakdown voltage, about 50 V, in order to make possible the operation of the active junction with high excess bias above the breakdown level. The devices were fabricated by a planar process in the laboratories of LAMEL-CNR, Bologna.

4. TESTS ON EPITAXIAL DEVICES AND CONCLUSIONS

The experimental tests confirm that the new devices provide a remarkably improved time resolution. Figure 4 shows the result of a test performed with a very fast diode laser. The

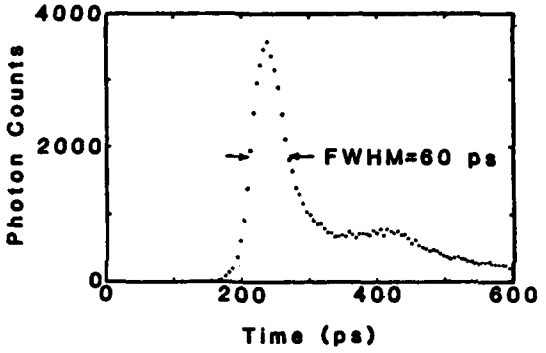


FIGURE 4

Resolution test on the p-p^{++} epitaxial SPAD, performed at 785 nm with a laser diode, producing a main pulse of about 40 ps duration followed by a small secondary pulse after 180 ps.

optical pulse width is specified to be about 40 ps, so that the FWHM resolution of the SPAD may be estimated to be 45ps or less. The dependence of the FWHM resolution of the new devices on the excess bias above the breakdown voltage was also measured. The results, reported in Figure 5, show that the behaviour is similar to that already observed for the previous non-epitaxial devices [9]. As illustrated in Figure 6, the experiments also showed that a remarkable reduction of the diffusion tail is obtained, in close agreement with the results of the Monte-carlo simulation.

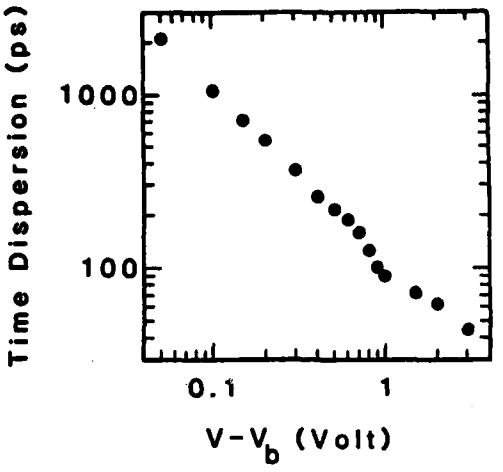


FIGURE 5

Dependence of the FWHM resolution of the p-p^{++} epitaxial SPAD device on the applied excess bias above the breakdown level.

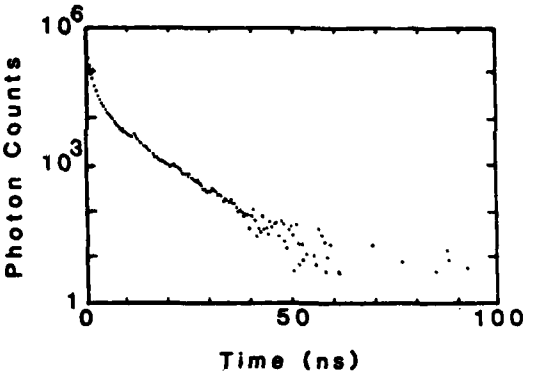


FIGURE 6

Resolution test on the p-p^{++} epitaxial SPAD device. The drawing scales and the experimental conditions are the same as in Figure 2; the fast peak in the time origin has 6.3 10^5 counts.

A natural question concerns the possibility of achieving further improvements in the time resolution. The situation is different for the fast peak and for the diffusion tail in the resolution function. The fast peak is dominated by the statistical physics of the avalanche build-up in the device; in fact, the other possible causes of time dispersion bring negli-

gible contributions. The carrier transit time dispersion is negligible, since the active junction depletion layer width is less than one micron; the trivial circuit jitter contribution was always verified to be quite smaller than the measured FWHM. Therefore, in order to ascertain how and to what extent the time resolution of the device can further be improved, a deeper and more detailed insight has to be gained in the statistical aspects of the avalanche build-up. Notwithstanding the remarkable amount of work on the avalanche statistics reported in the literature, this point deserves further attention, since it radically differs from other aspects that have found a satisfactory interpretation. In fact, the dependence on the electric field of the build-up time fluctuations is just opposite to the well known dependence of the multiplication noise in the amplifying mode, at bias lower than the breakdown voltage.

As concerns the diffusion tail, the physics of the effect is well understood and accurate computer simulation of different device structures can be obtained. With the epitaxial device structure here described, a further reduction of the diffusion tail implies using a p epitaxial layer with reduced thickness and a p^{++} substrate with shorter lifetime. However, a limit to the reduction of the epistrate thickness is set by the depth of the depletion layer of the n^- guard ring, necessary to obtain there a high breakdown voltage. The use of higher substrate doping does not appear to be advisable. In fact, the epitaxial devices are found to have dark counting rates of at least a few kHz, that is, significantly higher values than the previous non-epitaxial SPADs. This denotes that the lifetime in the active region of the new devices is shorter, and a substrate with higher doping is expected to introduce higher degradation in the epitaxial layer grown over it. It may be concluded that the further improvement obtainable with the $p\text{-}p^{++}$ epitaxial device structure is quite limited. Other device structures should therefore be investigated, in order to pursue a more drastic reduction of the tailing effect in the time resolution. Work is in progress in our laboratory on a device structure in p epitaxial layer over n substrate, in which the epistrate-substrate p-n junction acts as a sink for diffusing carriers.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. G. Ferla and his staff at SGS Microelettronica, Catania, for supplying the $p\text{-}p^{++}$ epitaxial wafers and gratefully acknowledge the active work of Dr. N. Camaioni and the collaboration of the technological laboratory of LAMEL-CNR, Bologna, in the fabrication of the devices.

REFERENCES

1. A. Goetzberger, B. Mc Donald, R. H. Haitz and R. M. Scarlett, J.Appl.Phys. 34, 1591 (1963).

2. R. H. Haitz, J.Appl.Phys. 36, 3123 (1965)

3. H. Sigmund, Infrared Phys. 8, 259 (1965)

4. H. Melchior, A. Goetzberger, E. Nicollian and W. T. Lynch, Solid State Electron. 12, 449 (1969)

5. R. J. Mc Intyre, IEEE Trans. Electron Devices ED-19, 703 (1972).

6. W.O. Oldham,R.R. Samuelson and P.Antognetti, IEEE Trans. Electron Devices ED-19, 1056 (1972).

7. S. Cova, A. Longoni and A. Andreoni, Rev. Sci Instrum. 52, 408 (1981).

8. S. Cova, A. Longoni. A. Andreoni and R. Cubeddu, IEEE J. Quantum Electron. QE-19, 630 (1983).

9. S. Cova, G. Ripamonti and A. Lacaita, Nucl. Instr. and Meth. A253, 482 (1987).

10. G. Ripamonti and S. Cova, Electron. Lett. 22, 818 (1985).

11. D.V. O'Conner and D. Phillips, Time-Correlated Single Photon Counting, Academic Press, London and New York (1983).

12. PPL30K Pulsed Diode Laser Modules, Opto-Electronics Inc., Oakville, Ontario.

13. J. Dziewor and W. Schmid, Appl. Phys. Lett. 31, 346 (1977).