SERIES IN MATERIALS SCIENCE AND ENGINEERING

# STRAINED-SI HETEROSTRUCTURE FIELD EFFECT DEVICES

C K MAITI
S CHATTOPADHYAY
L K BERA

Taylor & Francis
Taylor & Francis Group

# Strained-Si Heterostructure Field Effect Devices

## Series in Materials Science and Engineering

*Series Editors:* **Alwyn Eades,** Lehigh University, Bethlehem, PA, USA
**Evan Ma,** Johns Hopkins University, Baltimore, MD, USA

*Other books in the series:*

**Spintronic Materials and Technology**
Y B Xu, S M Thompson (Eds)

**Fundamentals of Fibre Reinforced Composite Materials**
A R Bunsell, J Renard

**Novel Nanocrystalline Alloys and Magnetic Nanomaterials**
B Cantor (Ed)

**3-D Nanoelectronic Computer Architecture and Implementation**
D Crawley, K Nikolic, M Forshaw (Eds)

**Computer Modelling of Heat and Fluid Flow in Materials Processing**
C P Hong

**High-K Gate Dielectrics**
M Houssa (Ed)

**Metal and Ceramic Matrix Composites**
B Cantor, F P E Dunne, I C Stone (Eds)

**High Pressure Surface Science and Engineering**
Y Gogotsi, V Domnich (Eds)

**Physical Methods for Materials Characterisation, Second Edition**
P E J Flewitt, R K Wild

**Topics in the Theory of Solid Materials**
J M Vail

**Solidification and Casting**
B Cantor, K O'Reilly (Eds)

**Fundamentals of Ceramics**
M W Barsoum

**Aerospace Materials**
B Cantor, H Assender, P Grant (Eds)

**Physical Methods for Materials Characterisation**
P E J Flewitt, R K Wild

Series in Materials Science and Engineering

# Strained-Si Heterostructure Field Effect Devices

**C K Maiti**
*Indian Institute of Technology, India*

**S Chattopadhyay**
*University of Newcastle, UK*

**L K Bera**
*Kulim Hi-Tech, Kedah, Malaysia*

Taylor & Francis
Taylor & Francis Group
New York London

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2007 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 0-7503-0993-8 (Hardcover)
International Standard Book Number-13: 978-0-7503-0993-6 (Hardcover)

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

No part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC) 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

**Library of Congress Cataloging-in-Publication Data**

Maiti, C. K.
Strained-Si heterostructure field effect devices / C.K. Maiti, L.K. Bera, and S. Chattopadhyay.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-0-7503-0993-6 (alk. paper)
1. Metal oxide semiconductor field-effect transistors. 2. Silicon--Electrical properties. I. Bera, L. K. II. Chattopadhyay, Swapan, 1952- III. Title. IV. Title: Strained silicon heterostructure field effect devices.

TK7871.95.M27 2007
621.3815'284--dc22 2006031518

**Visit the Taylor & Francis Web site at**
**http://www.taylorandfrancis.com**

**and the CRC Press Web site at**
**http://www.crcpress.com**

# *Dedication*

WHERE ARE WE HEADING?

The principal applications of any
sufficiently new and innovative technology
always have been
- and will continue to be -
applications created by that new technology

**..... Herbert Kroemer**

**We dedicate this Monograph**
**Strained-Si Heterostructure Field Effect Devices**
**to YOU, Sir.**

# *Preface*

Soon after Bardeen, Brattain, and Shockley invented the solid-state device in 1947, the microelectronics revolution started. Today, the world market for semiconductors is beyond 200 billion dollars, of which 80% is dominated by CMOS and interestingly, Si devices account for 97% of all microelectronics. The semiconductor sales have doubled in every three technology generations. However, the semiconductor industry is now facing tremendous challenges to extend the integrated circuit (IC) performance to meet the ever-increasing demand for complex functionality. Innovations in both the device structures and new materials are needed to continue with the Si CMOS roadmap.

While introducing new structure innovations has always been an important part of device scaling, the integration of new materials is facing serious hurdles in order to meet the aggressive specifications of the International Technology Roadmap for Semiconductors (ITRS). Although the idea of using silicon-germanium (SiGe) and silicon strained-layer epitaxy toward bandgap engineering in semiconductor devices is an old one, this concept has become a reality only during the last two decades. It is expected that strained-Si will set a new technology roadmap for the CMOS industry with enhanced performance, however, with a minimum change in the existing Si infrastructure. This is a solution the industry has been searching for the last few years.

Since the mid-nineties, there is a growing interest in bipolar and MOSFETs designed with SiGe and strained-Si/SiGe as channels because of the potential enhancement in performance due to improved carrier transport properties. Recent integration of III-V layers on Si substrates has created tremendous interest, even though the direct application of such technologies may not take place in the near future. Novel channel materials may also support operation at high power or temperature and lead to increased radiation hardness, and band structure engineering would offer design flexibility and monolithic optoelectronic circuits would likely become manufacturable. While the new devices are expected to be highly innovative, they still have to meet the fundamental criteria of functional capability and economic viability.

The purpose of the monograph is to survey the R&D strategy and address the research needs associated with the front-end aspects of extending CMOS technology via strain engineering. Areas covered include transistor structures, front-end materials and processes along with their characterization, modeling and simulation. Although lithography, back-end processes, circuit architecture, design and test are very important areas that need to be addressed for extending the CMOS, they are beyond the scope of this monograph.

As the new device structures are expected to provide advanced features from the materials in current use, the "material limit" is also being reached. Additional performance boosts can only be expected from new materials with superior electronic properties. Such properties may include high mobility (in channel), high dielectric constant (gate dielectric), controllable work functions (gate electrode), or low-resistivity (source/drain and associated contacts). The integration challenges are far more complex than those involved in the fabrication of bulk materials with specified properties.

Epitaxial technologies continue to emerge as a dominant processing step for Si-based devices. The process development for strained-Si on SiGe further points towards epitaxial processes playing an increasingly important role in future electronic and optoelectronic devices. Introduction of new materials brings in new metrology needs. These include metrologies for high spatial-resolution strain measurement, atomic force microscopy for surface analysis, dopant profiling and dose monitoring in small area and nonplanar topographies, and characterization of surface mobility in new substrates, such as, SiGe and strained-Si. Similar requirements are in the modeling and simulation, with the understanding that the new modeling and simulation tools have to be provided with the supporting database and model parameters for the materials involved.

Extension of CMOS to 22-nm node and beyond may require new nonclassical MOSFET structures coupled with advanced materials and processes. Classes of new materials include high-k gate dielectrics, metal and midgap gate metal electrodes, strained-Si and silicon-germanium alloys. These new materials will lower the gate leakage current and gate resistance and reduce the poly-gate electrode depletion capacitance, and increase the device speed. Nonclassical CMOS structures offer better control of short-channel effects, improved $I_{on}$ via higher channel mobility, lower load capacitance and lower propagation delay time. These structures also offer lower $I_{off}$ and lower switching energy.

Currently, strained-Si is the performance booster for extending Moore's law. Local strain (process-induced strain) approach towards strained-Si engineering, perhaps would be the simplest among the possible approaches. However, scaling issues of local strain approach may require wafer-level uniaxial strain to provide most of the total strain for future device generations. It is now believed that wafer scale (substrate-induced) strained-Si will be as effective as silicon-on-insulator (SOI) technology in terms of extending CMOS performance. Additionally, the combination of new device structures and new materials enables new operating principles that may provide new behavior and functionality beyond the constraints of bulk planar or classical CMOS. Implementation into manufacturing of nonclassical MOSFET device structures integrated with new materials and processes (for example, a strained-Si channel integrated with a new high-k gate dielectric material) will be interesting.

The theoretical assessment via simulation of nanodevices is a challenge be-

cause band structures, arbitrary wafer orientation, quantum effects, and electrostatics all are needed to be considered simultaneously. A systematic analysis of the strain effects on bandstructure and mobility in deeply scaled n- and p-MOSFETs with Si, SiGe, strained-Si and Ge channel is essential. In order to assess and compare the performances of nanoscale CMOS devices with novel channel materials (e.g., SiGe, strained-Si, and Ge), a new simulation approach is also necessary.

In this book, current status and future trends in the strained-Si materials and devices are discussed. This book is devoted to research and development (R&D) Engineers and Scientists who are actively involved in extending the roadmap for Si CMOS devices mainly via strain engineering. This monograph may also serve as a reference for postgraduate and research students. By introducing a diversity of R&D activities and results in this book, we hope to ignite even more novel concepts and innovative thinking in the field.

**C K Maiti**
**Indian Institute of Technology, Kharagpur, India**

**S Chattopadhyay**
**University of Newcastle, UK**
**and**
**L K Bera**
**Kulim Hi-Tech, Kedah, Malaysia**

# Contents

**1 Introduction** **1**
1.1 Heterostructure Field-Effect Devices . . . . . . . . . . . . . . . 7
1.2 Substrate Engineering . . . . . . . . . . . . . . . . . . . . . . . 8
1.2.1 Substrate-Induced Strain Engineering . . . . . . . . . 9
1.2.2 Process-Induced Stress Engineering . . . . . . . . . . . 10
1.2.3 Orientation-Dependent Mobility Engineering . . . . . 11
1.3 Gate Dielectrics on Engineered Substrates . . . . . . . . . . . 12
1.4 Strained-Si Technology: Process Integration . . . . . . . . . . 13
1.5 Nonclassical CMOS Structures . . . . . . . . . . . . . . . . . . 14
1.6 Strain-Engineered Hetero-FETs: Modeling and Simulation . . 16
1.7 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

**References** **18**

**2 Strain Engineering in Microelectronics** **23**
2.1 Stress Induced During Manufacturing . . . . . . . . . . . . . . 26
2.2 Global vs. Local Strain . . . . . . . . . . . . . . . . . . . . . . 32
2.3 Substrate-Induced Strain . . . . . . . . . . . . . . . . . . . . . 35
2.4 Process-Induced Stress . . . . . . . . . . . . . . . . . . . . . . 37
2.4.1 Mechanical Stress Control . . . . . . . . . . . . . . . . 43
2.4.2 Substrate Orientation Dependence . . . . . . . . . . . 46
2.4.3 Strained-Ge . . . . . . . . . . . . . . . . . . . . . . . . 48
2.5 Stress/Strain Analysis . . . . . . . . . . . . . . . . . . . . . . . 48
2.6 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 52

**References** **53**

**3 Strain-Engineered Substrates** **59**
3.1 Epitaxy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60
3.2 Heteroepitaxy and Strain Control . . . . . . . . . . . . . . . . 64
3.3 Engineered Substrates: Technology . . . . . . . . . . . . . . . 71
3.3.1 Virtual Substrates . . . . . . . . . . . . . . . . . . . . . 73
3.3.2 Substrate Specifications . . . . . . . . . . . . . . . . . 79
3.3.3 Strained-Si on Insulator . . . . . . . . . . . . . . . . . 81
3.3.4 Smart-Cut . . . . . . . . . . . . . . . . . . . . . . . . . 87
3.3.5 Hybrid Substrates . . . . . . . . . . . . . . . . . . . . . 89
3.3.6 GeOI Substrates . . . . . . . . . . . . . . . . . . . . . . 90

3.4 Characterization of Strained Layers . . . . . . . . . . . . . . . . 93
3.4.1 XRD . . . . . . . . . . . . . . . . . . . . . . . . 96
3.4.2 SIMS . . . . . . . . . . . . . . . . . . . . . . . . 100
3.4.3 Raman . . . . . . . . . . . . . . . . . . . . . . . . 101
3.5 Engineered Substrates . . . . . . . . . . . . . . . . . . . . 106
3.5.1 Mobility Comparison . . . . . . . . . . . . . . . . . . . . 106
3.5.2 Thermal Conductivity . . . . . . . . . . . . . . . . . . . . 109
3.6 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . 111

**References** **111**

**4 Electronic Properties of Engineered Substrates** **119**
4.1 Substrate-induced Strained-Si . . . . . . . . . . . . . . . . . . . . 120
4.1.1 Energy Gap and Band Structure . . . . . . . . . . . . . 122
4.1.2 Electron Mobility . . . . . . . . . . . . . . . . . . . . 134
4.1.3 Hole Mobility . . . . . . . . . . . . . . . . . . . . . . 138
4.1.4 Field Dependence . . . . . . . . . . . . . . . . . . . . 141
4.1.5 Doping Dependence . . . . . . . . . . . . . . . . . . . . 145
4.2 Carrier Lifetime . . . . . . . . . . . . . . . . . . . . . . . . 146
4.3 Mobility: Thickness Dependence . . . . . . . . . . . . . . . . 150
4.4 Mobility: Temperature Dependence . . . . . . . . . . . . . . . . 157
4.5 Diffusion in Strained-Si . . . . . . . . . . . . . . . . . . . . 159
4.6 Process-induced Strained-Si . . . . . . . . . . . . . . . . . . . . 160
4.6.1 Hole Mobility . . . . . . . . . . . . . . . . . . . . . . 166
4.6.2 Electron Mobility . . . . . . . . . . . . . . . . . . . . 170
4.7 Uniaxial vs. Biaxial Strain Engineering . . . . . . . . . . . . . 176
4.8 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . 179

**References** **180**

**5 Gate Dielectrics on Engineered Substrates** **189**
5.1 Strained-Si MOSFET Structures . . . . . . . . . . . . . . . . . 190
5.2 Thermal Oxidation of Strained-Si . . . . . . . . . . . . . . . . 191
5.2.1 Ge Diffusion . . . . . . . . . . . . . . . . . . . . . . . 193
5.2.2 Kinetics: Oxidation of $Si_{1-x}Ge_x$ Layers . . . . . . . . 197
5.2.3 Oxidation of Strained-Si Layers . . . . . . . . . . . . . 198
5.3 Rapid Thermal Oxidation . . . . . . . . . . . . . . . . . . . . 200
5.4 Plasma Nitridation of Strained-Si . . . . . . . . . . . . . . . . 203
5.5 Effect of Surface Roughness . . . . . . . . . . . . . . . . . . . 207
5.6 Effect of Strained-Si Layer Thickness . . . . . . . . . . . . . . 210
5.7 High-k Gate Dielectrics on Strained-Si . . . . . . . . . . . . . 220
5.7.1 Microwave Plasma Deposition . . . . . . . . . . . . . . . 225
5.7.2 Chemical Analysis . . . . . . . . . . . . . . . . . . . . 226
5.7.3 Conduction Mechanism . . . . . . . . . . . . . . . . . . 230
5.7.4 Reliability Issues . . . . . . . . . . . . . . . . . . . . 232

5.8 Gate Dielectrics on Ge . . . . . 234
5.9 Summary . . . . . 234

**References** **235**

**6 Heterostructure SiGe/SiGeC MOSFETs** **245**
6.1 SiGe/SiGeC: Material Parameters . . . . . 247
6.2 SiGe Hetero-FETs: Structures and Operation . . . . . 250
6.3 SiGe p-MOSFETs on SOI . . . . . 267
6.4 SiGeC Hetero-FETs . . . . . 273
6.5 SiGe-based HEMTs . . . . . 277
6.6 Design Issues . . . . . 281
6.6.1 Gate Engineering . . . . . 281
6.6.2 Layer Design . . . . . 286
6.7 Summary . . . . . 287

**References** **288**

**7 Strained-Si Heterostructure MOSFETs** **295**
7.1 Operating Principle . . . . . 296
7.1.1 Threshold Voltage . . . . . 297
7.2 Uniaxial Stress: Process Flow . . . . . 306
7.3 Strained-Si MOSFETs with SiC-Stressor . . . . . 309
7.4 Biaxial Strain: Process Flow . . . . . 315
7.5 Scaling of Strained-Si MOSFETs . . . . . 323
7.5.1 Layout Dependence . . . . . 333
7.5.2 Thickness Dependence . . . . . 338
7.5.3 Orientation Dependence . . . . . 341
7.5.4 Hetero-FETs: Single Gate vs. Double Gate . . . . . 344
7.5.5 Hetero-FET: Dual Channel Structure . . . . . 346
7.6 Strained-Si MOSFETs: Reliability . . . . . 353
7.6.1 Self-heating . . . . . 358
7.7 Industry Example: TSMC . . . . . 360
7.8 Industry Example: AMD . . . . . 370
7.9 Summary . . . . . 371

**References** **376**

**8 Modeling and Simulation of Hetero-FETs** **385**
8.1 Simulation of Hetero-FETs . . . . . 386
8.2 Strained-Si Material Parameters for Modeling . . . . . 387
8.3 Simulation of Strained-Si n-MOSFETs . . . . . 390
8.4 Characterization of Strained-Si Hetero-FETs . . . . . 399
8.5 TCAD: Strain-engineered Hetero-FETs . . . . . 400
8.6 SPICE Parameter Extraction . . . . . 409

8.7 Performance Assessment . . . . . . . . . . . . . . . . . . . . . . . . 412
8.8 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 415

**References** **417**

**Index** **421**

# 1

# *Introduction*

Technology changes along the lines of materials; we had the Stone Age, the Bronze Age and the Iron Age. Maybe the last century will go down in history as the Silicon Age. The field of material sciences has always provided the innovations for new technology. Silicon has been the workhorse of the semiconductor industry over the past 50 years, the majority of the improvements in devices has been achieved via device scaling.

For the field-effect devices, as described in two patents of Lilienfeld [1, 2] and in a later patent by Heil [3], it took 35 years to turn their ideas into a usable structure where a substantial modulation of the conduction could be achieved. To reach this point fundamental insight into the characteristics of semiconductor had to be obtained. In the forties and fifties of the last century, the development of the semiconductor technology was directed toward bipolar devices, whereas field effect devices received hardly any attention. The first working metal oxide semiconductor field-effect transistor (MOSFET) became available in 1960 [4]. This foundation let the MOSFET technique live up to its potential and developed over the next 40 years into a technology of great importance for the fabrication of ever faster devices of ever-smaller dimensions at an ever-lower price.

Fig. 1.1 schematically illustrates the basic MOSFET device used in today's silicon chips. The basic fabrication process steps to manufacture such a device have been broadly described in reference [5]. There are no serious competitors to replace this device in the foreseeable future. The basic structure will continue to evolve to allow continued performance improvements, but fundamental changes are unlikely in the next 10 years. The evolution of Si semiconductor technology has so far followed a trend described by Moore's law, an industry axiom that predicts that the number of transistors on a chip will double every 18 months, largely due to continued miniaturization known as scaling. Scaling involves shrinking all of the critical dimensions of the device (such as the gate or channel length and gate oxide thickness) by roughly the same factor. This in turn reduces the length of the path that electrons and holes have to travel, making the current in the devices flow faster and consume less power. Scaling is becoming more challenging with each successive technology node, as fundamental technological limits of existing processes and materials are being reached. The huge costs of scaling CMOS devices have left the Si industry at a crossroad as Si chips can no longer shrink in size.

Silicon has so far served us well, but its fundamental properties in terms

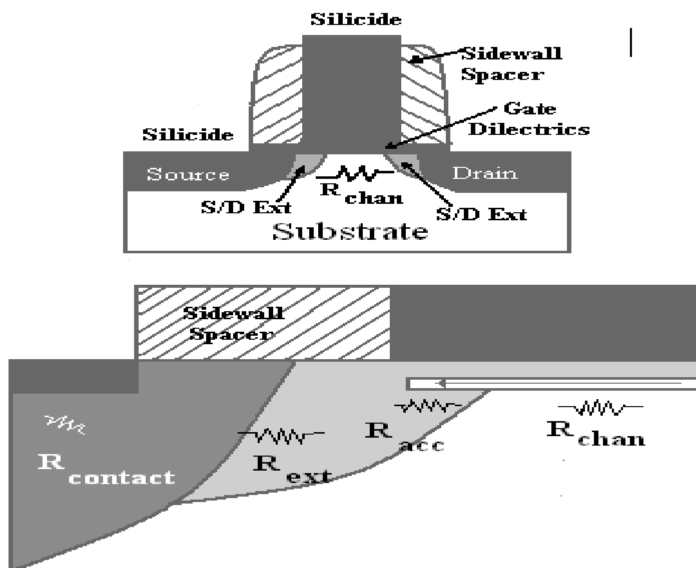



**FIGURE 1.1**
Schematic cross-section of a modern MOS transistor. After J. D. Plummer and P. B. Griffin, *Proc. IEEE*, Vol. 89, 2001(240-258). With permission.

of carrier mobility and associated intrinsic performance are limited. These limitations are now being exposed for various applications and the base Si substrate material is being replaced for many devices. Also, continued transistor scaling will not be as straightforward in the future, as fundamental limits are being reached not only in some of the key materials and processes, but most importantly in the power density of these ultrasmall devices. With gate lengths below 50-nm and gate oxides only a few atomic layers thick in modern technologies, the amount of leakage current, in particular, is beginning to dominate device behavior and limiting the capabilities of the conventional downscaling. The major technology barriers to scaling CMOS and improving performance now are: scaling power supply voltage, sustaining low leakage current, and increasing drive current.

As scaling of Si CMOS is approaching its limits, innovative device structures and materials are actively being investigated to boost the performance. Manufacturing practices in the Si CMOS industry are currently experiencing difficulty, where new approaches are necessary in order to maintain improved product capability within reasonable economic constraints. When approaching the 22-nm technology node, advanced device structures will be needed to overcome fluctuation problems associated with the small number of dopants in the active region. While the new devices are expected to be highly innovative, they still have to satisfy the fundamental criteria of functional scalability

**TABLE 1.1**
Evolution of CMOS technology.

| Technology node $\mu$m | Year | Metal level | Supply voltage (V) | Oxide thickness (A) | $V_t$ (V) |
|---|---|---|---|---|---|
| 0.7 | 1988 | 2 | 5.0 | 200 | 0.7 |
| 0.5 | 1992 | 3 | 3.3 | 120 | 0.6 |
| 0.35 | 1994 | 5 | 3.3 | 75 | 0.5 |
| 0.25 | 1996 | 6 | 2.5 | 65 | 0.45 |
| 0.18 | 1999 | 7 | 1.9 | 50 | 0.40 |
| 0.12 | 2001 | 8 | 1.5 | 40 | 0.30 |
| 90-nm | 2004 | 8-9 | 1.0 | 35 | 0.25 |

and economic viability.

The improved transport properties of new channel materials, such as strained-Si, SiGe, Ge and III-V semiconductors, along with new device designs, such as dual-gate, tri-gate or FinFETs, are expected to enhance the performance of nanoscale CMOS devices. Novel process techniques, such as atomic layer deposition (ALD), high-k dielectrics, and metal gates are now being used in exploratory device structures. New materials in the channel promise reduced series resistance and higher drive currents. As the end of the Si roadmap is approached, strained-Si channel will offer a way to improve the MOSFET drive current along with transistor downscaling.

Current nanometer-scaled MOSFETs are true short-channel devices, the structure of which has to withstand drain-induced barrier lowering (DIBL) without exceeding the maximum source-drain leakage current that can be tolerated by the application. Devices dimensions are now well into tens of nanometer regime and serious doubts are being raised regarding the ability to scale down the gate length of conventional bulk-Si MOSFET below 30 nm. Because the subthreshold slope and built-in potentials cannot be scaled, the threshold voltage cannot be reduced too much without the leakage current becoming excessive. Although the total power dissipation could be managed by forced cooling, the power dissipation through drain-to-source and gate-to-channel leakage currents will become increasingly detrimental to circuit operation and will limit overall performance.

Table 1.1 shows a set of key MOSFET parameters and their evolution with time and the technology node. It is important to note that there has been a tremendous increase in number of metal interconnects leading to increased circuit density and reduction in power supply voltage with lower power consumption. Reduction in the gate oxide thickness down to atomic scale values, although increasing the speed, leads to higher leakage current. The power supply decrease is the consequence of the reduction of the MOS gate oxide thickness, which, however, induces a lower oxide breakdown voltage. One also notices the slow decrease of the threshold voltage of the MOS devices. The results indicate that even if highly idealized device structures are assumed,

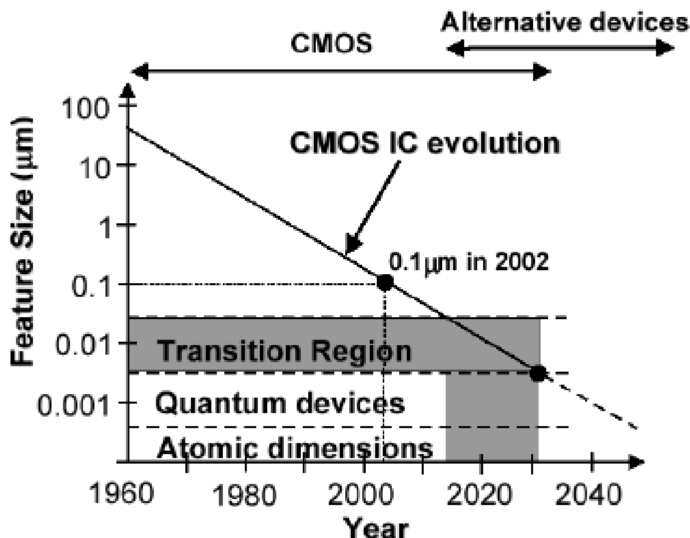



**FIGURE 1.2**
Feature size vs. time in silicon ICs. After J. D. Plummer and P. B. Griffin, *Proc. IEEE*, Vol. 89, 2001(240-258). With permission.

only limited improvements over the conventional MOSFETs can be achieved by the channel engineering techniques. Most of the history represented in Fig. 1.2 has been achieved with MOS transistors, and with a limited number of materials such as Si, $SiO_2$, Al, $Si_3N_4$, $TiSi_2$, TiN, and W. Fig. 1.2 summarizes the trends in feature size over time.

Recently a 65-nm generation logic technology with 1.2-nm physical gate oxide, 35-nm gate length, enhanced channel strain, NiSi, 8 layers of Cu interconnect, and low-k interlayer dielectric (ILD) for dense high performance logic has been reported [6]. Transistor gate length is scaled down to 35-nm while not scaling the gate oxide as a means to improve performance and reduce power. Increased nMOS and pMOS drive currents are achieved by enhanced channel strain and junction engineering. The supply voltage cannot be scaled as fast as the dimensions of the devices, resulting in larger electric fields and reliability issues. Because the subthreshold slope and built-in potentials cannot be scaled, the threshold voltage cannot be reduced too much without the leakage current becoming excessive.

Higher channel dopant concentrations and more abrupt, shallower source-drain (SD) junctions have been used to control short-channel effects (SCE) at very short gate lengths. Doping requires quite precise profile design and process control, whereas increasing channel doping concentration degrades carrier mobility, lowering the drain current. Statistical fluctuation of channel

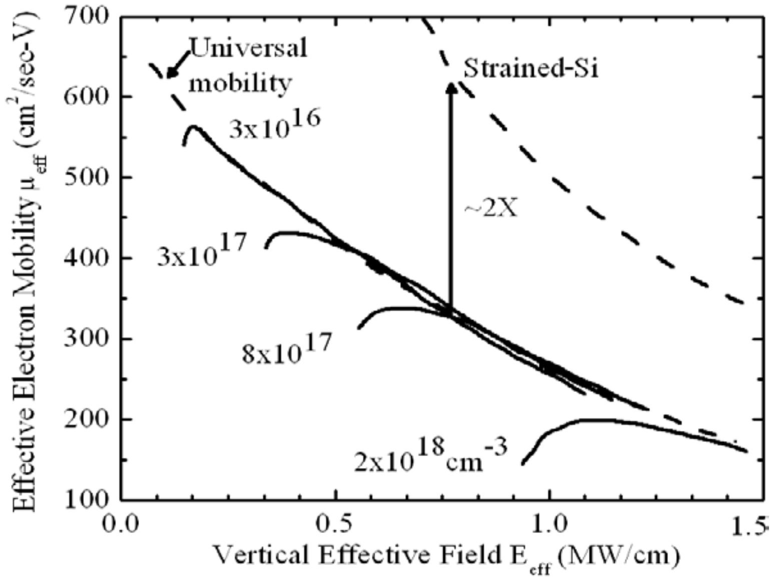



**FIGURE 1.3**
Channel transport and universal mobility in Si. Electron mobility may be enhanced ($\sim 2\times$) by substrate-induced strain in Si. After S.-I. Takagi, ULIS 2004 Lecture Materials.

dopants causes increasing variation of the threshold voltage, posing difficulty in circuit design while scaling the supply voltage. These result in detrimental effects such as degraded mobility, dopant fluctuation effects, and increased series resistance.

In order to circumvent some of the scaling issues, new materials and device architectures are being integrated into CMOS devices in order to maintain the historic CMOS performance trend. The important research needs may be organized into six major categories;

(a) Substrate Engineering,
(b) Gate Stack,
(c) Junctions and Contacts,
(d) CMOS Compatible Device Structures,
(e) Technology Modeling and Simulation, and
(f) Characterization.

Given the limitations associated with further reduction of the gate length in the quest to continue to increase device speed, the only alternative is to enhance the carrier (electrons and holes) mobility in the channel. Mobility can be improved by using a variety of new Si technologies, including strained silicon (strained-Si), silicon-germanium (SiGe), SiGe on silicon-on-insulator, as well as combinations of these materials (see Fig. 1.3). The impact of

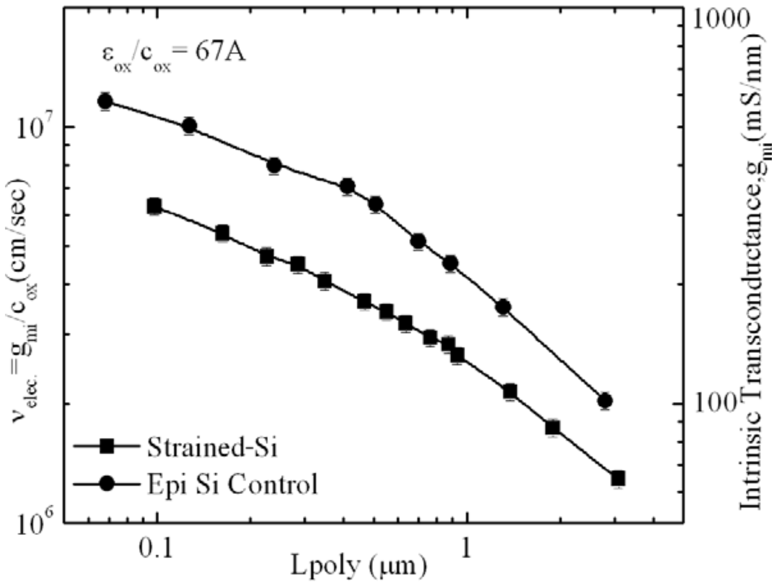



**FIGURE 1.4**
Biaxial strain increases electron mobility above the universal Si mobility curve. Both $I_d$ and $g_m$ are enhanced. After K. Rim et al., *IEEE IEDM Tech. Dig.*, 1998(707-710). With permission.

enhanced mobility on drive current in a strained-Si n-MOSFET is shown in Fig. 1.4. One of the most promising alternatives for the replacement of bulk-Si substrates in the CMOS industry is now the heterostructure Silicon technology. For the Si substrate, silicon-on-insulator, which improves chip performance via capacitance reduction, has been introduced in production and is on the roadmap of most major Si chip companies. Faster chips can be made by using other materials such as gallium arsenide. But gallium arsenide is much more expensive than Si. Worse yet, making chips of it requires an entirely different process than that used for Si.

Increasingly new materials need to be introduced in technology development due to physical limits which otherwise would prevent further scaling. This is required especially for gate stacks, interconnect structures, and photoresists. In consequence, equipment, process, device, and circuit models must be extended to include these new materials. Furthermore, computational material science needs to be developed and applied to contribute to the assessment and selection of new materials in order to reduce the experimental effort. An excellent review of CMOS material and structural changes during the past 40 years, and their future trends, can be found in reference [7].

In this book, we will focus primarily on "front-end" issues (structures, processes, and materials associated with the switching devices). However, the objective of this chapter is to highlight some of these new innovations, and to

survey the current progress and potential. We shall focus on front-end materials and processes, as well as advanced transistor structures. It is particularly intended to address the research needs associated with the front-end aspects of extending CMOS technology.

Areas including transistor structures, front-end materials and front-end processes, along with their characterization and modeling/simulation infrastructure to extend front-end fabrication of CMOS from the 65-nm node to its ultimate limit at or beyond the 22-nm node, will be covered. Research needs related to "beyond CMOS" technologies will also be addressed. Historically, new process generations have been limited by the lithographic capability to transfer ultra-small patterns required by the new technology. However, the needs associated with lithography, back-end processing, design, and test are beyond the scope of this book.

## 1.1 Heterostructure Field-Effect Devices

Although the SiGe HBTs at present dominate the commercial silicon heterostructure devices, the first Si-based heterostucture field-effect transistor (hetero-FET) was demonstrated in 1986 [8]. These earliest hetero-FETs were Schottky-gated, III-V-like n- and p-channel modulation doped devices, which gave rise to a variety SiGe-based MODFET topologies. Recently, the field-effect devices have centered on strained-Si MOSFETs, because of their better compatibility with mainstream CMOS processing, and the impressive mobility enhancements that have been realized in aggressively scaled gate lengths. Two fundamentally different ways of producing strained-Si CMOS exist, utilizing both biaxial and uniaxial strain techniques.

Strained-Si technology, the introduction of elastic strain in Si transistor channels, is the most widely accepted method for enhancing the carrier mobility of Si. There are two variants to strained-Si technology that have been researched and demonstrated: (a) local (process-induced) strain introduction via transistor module engineering and (b) global (wafer-scale) strain introduction via SiGe (SiGe) epitaxial processes (and layer transfer processes for global strain with SOI). Besides using substrate to provide stress in channel, strain effects induced during the fabrication process can increase the channel mobility. Both tensile and compressive stresses can be introduced to the channel in any one of the three dimensions by different process techniques. Several stress-transfer techniques (such as etch-stop liner and stress transfer via strained-SiGe) using extra integration process into an existing baseline process have been reported. In addition, new preparation techniques of strained-Si surface (e.g., biaxial tensile stress) and different substrate orientation to enhance mobility will be introduced. The challenges and importance of

**TABLE 1.2**
Reported enhancements in MOSFETs with local and global strain engineering techniques.

| Manufacturer | Strain Type | $L_g$ (nm) | nMOS $I_{dsat}$ Enhancement | pMOS $I_{dsat}$ Enhancement | Ref. |
|---|---|---|---|---|---|
| Freescale/TSMC/ SOITEC | Global (on SOI) | 150 | 28% | 12% | [11] |
| IBM | Global | 67 | 35% | N/A | [12] |
| IBM/Sony/Toshiba AMD/Chartered | Local (on SOI) | 45 | 11% | 20% | [13] |
| Intel | Local | 45 | 10% | 25% | [14] |
| Intel | Global | 140 | 24% | N/A | [15] |
| Toshiba | Global | 40 | N/A | 19% | [16] |
| AMD | Local (on SOI) | 40 | 13% | 20% | [17] |
| AMD | Global | 25 | 25% | N/A | [18] |
| TSMC | Global | 60 | 15% | N/A | [19] |

each method will be discussed and compared. In addition, we shall highlight how the stress oriented from the layout geometry affects the device electrical behavior. The issues and improvement in the circuit level device modeling will be discussed.

Channel engineering in the form of strained-Si on relaxed-SiGe substrate, compressive SiGe in the channel or in recessed SD region has received much attention as materials options to increase drive current in CMOS devices [9]. However, there exists many challenges such as dislocations, surface roughness, Ge and dopant diffusion which need to be addressed for successful integration of strained-Si in mainstream CMOS technology. High-k gate dielectrics offer another opportunity to enhance CMOS devices by enabling scaling the effective oxide thickness (EOT) of gate dielectrics to below 1-nm with significantly reduction of gate leakage currents. However, the implementation of high-k gate dielectrics is still problematic in CMOS devices. The use of high-k dielectrics can result both in mobility and reliability degradation [10]. Table 1.2 summarizes several transistor performance results based upon global and local strain engineering techniques.

## 1.2 Substrate Engineering

Introduction of strain to delay the need for integration of new materials, such as high-k and further channel/gate scaling by at least one generation, may be viewed as a viable technology node extender. Mobility enhancement in Si has

been demonstrated utilizing engineered substrates, for example, by straining the Si channel. Electron and hole mobility may also be optimized separately by patterning substrates to expose different Si crystal faces for each device. Even faster complementary MOS (CMOS) logic operation would be attainable with high mobility channel materials, e.g., Ge or III-V semiconductors. New metrology needs for material and structure characterization are rising. These include metrologies for high spatial-resolution strain measurement of patterned surfaces, dopant profiling and dose monitoring in small area and nonplanar topographies, and characterization of surface mobility in new substrates such as SiGe and strained-Si.

### 1.2.1 Substrate-Induced Strain Engineering

Strained-Si technology enables improvements in CMOS performance and functionality via replacement of the bulk, crystalline-Si substrate with a Si substrate that contains a tetragonally distorted, biaxially strained-Si thin film at the surface. In fact, the improved performance comes from the properties of the strained-Si itself. Rather than shrinking the size of the transistor, strained-Si technology changes the nature of the wafer itself by slightly stretching the placement of the atoms. This increases mobility permitting electrons to travel faster. In strained-Si technology, the substrate fabrication involves generally UHV epitaxy techniques for large-scale production. However, strained-Si layers have also been produced using industry-standard low-pressure Si epitaxy systems, such as the ASM Epsilon and Applied Materials Epi Centura platforms.

The strained-Si layer is produced by depositing a thin layer of Si on an alloy layer of Si and Ge onto a Si substrate. The top Si layer is strained at the Si/SiGe interface because lattice of SiGe exerts a strain on the thinner top Si layer, stretching the Si lattice slightly. Further, by controlling the amount of Ge in the bottom SiGe layer, the amount of strain produced in the overlying Si layer can be manipulated. It was discovered that the lattice mismatch between Si and Ge atoms could be accommodated by a finite degree of lattice distortion [20]. This distortion or strain actually offers the advantage of allowing electrons and holes to move faster (or have higher mobility). This increase in carrier mobility is attributed to a modified Si band structure that lowers the resistance to electron and hole movement in the material. As a result, the use of strained-Si increases electron and hole mobility, and results in higher drive current (10-35%). Increased drive current means increased transistor speed and a 30% speed gain can be translated into roughly two-times lower power consumption. Thus, strained-Si offers a relatively easy way to make more powerful chips with low-power consumption.

Besides strained-Si in the traditional (100) plane, it may be advantageous to change the crystal orientation to optimize CMOS circuit performance. The carrier mobility of inversion layers depends on surface orientation and current flow directions, due to asymmetry of the carrier effective masses in the Si

crystal lattice [21]. Recent progress has also demonstrated the evolution of the strained-Si bulk MOS structure, such as the strained-Si on SiGe on insulator (SGOI) MOSFET [22, 23], and the strained-Si directly on insulator (SSDOI) MOSFET [24, 25, 26]. Having a highly strained-Si channel [27] or using a different orientation (110) substrate [28] in p-type MOSFETs, the performance match between the n- and p-MOSFET for CMOS applications might be achieved.

Strained-Si holds out the promise of a big performance enhancement without the need for a major overhaul of the existing chip manufacturing process. However, strained-Si wafers cost more, while a standard eight-inch wafer is about $120, the strained-Si wafer may cost 2 to 3 times more. Growing the kinds of SiGe-graded layers required to produce strained-Si typically poses two problems: (a) defect propagation into the strained-Si layer and (b) surface crosshatch. Chemical-mechanical polishing in conjunction with epitaxy techniques developed by Amberwave, makes grading of the epitaxial layer for producing good strained-Si layer.

Freescale Semiconductor and the Soitec have made an important advancement in strained-Si technology for next-generation devices. They have announced an approximate 70% increase in electron mobility, as well as high compatibility with existing SOI CMOS processes while optimizing CMOS device performance at the sub-65 nm nodes using strained silicon-on-insulator (SSOI) engineered substrates.

### 1.2.2 Process-Induced Stress Engineering

Process-induced stress requires different stress types (compressive and tensile for n- and p-channel, respectively) to simultaneously improve both n- and p-channel devices. For the last several years, strain developed during semiconductor manufacturing was becoming recognized as offering the best potential to enhance performance in sub-100 nm process technologies. A novel strained transistor architecture has been developed, incorporating epitaxially grown strained-SiGe source-drain regions for p-MOSFETs and a highly tensile silicon nitride capping layer for n-MOSFETs [14]. Although process-induced stress has been used to achieve significant mobility enhancements in short-channel devices (e.g., $\sim 2\times$ for holes in 65-nm technology), the Si/SiGe materials system has the potential to achieve very large improvements in mobility.

Intel's approach to form strained-Si through a MOS transistor channel is by using embedded SiGe, which attributed to uniaxial compressive channel stress. Intel's announcement that it would put strained-Si into production in its 90-nm process for forthcoming Pentium 4 processors has underlined the growing gap between the technologies. The integrated Intel process combines strained-Si processing, copper interconnects, and low-k dielectric materials with high-performance, and low-power transistors. The transistors in Intel's 90-nm process have gate lengths of 50-nm and have the gate oxide thickness of 1.2 nm. Thinner gate oxides and ultra-shallow doping profiles help make

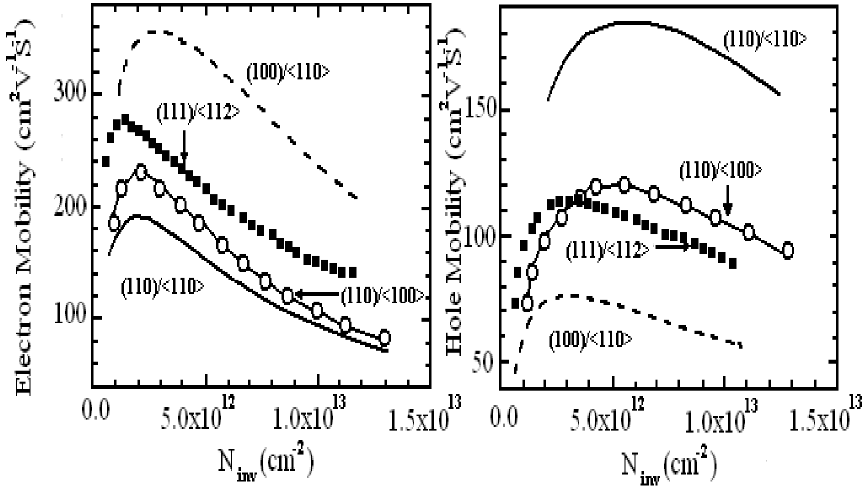



**FIGURE 1.5**
Carrier mobility in the inversion layer with various substrate orientations. Electron mobility is highest on (100) substrate while hole mobility is highest on (110) substrate with current flow along <110> direction. The gate oxides (about 2-nm thick) were grown simultaneously in $N_2O$ ambient at $800^oC$ in conventional thermal furnace. After M. Yang et al., *IEEE IEDM Tech Dig.*, 2003(453-456). With permission.

sure the transistor works properly with such a short channel length.

Although putting a controlled layer of Ge into place is a challenging task, the improvement in the current drive and, therefore, the switching speed of transistors is expected to outweigh the extra cost involved in using the strained-Si technology. However, in this approach, it is difficult to predict how much advantage may be achieved as one has to overcome the surface defects that can interfere with gate formation.

### 1.2.3 Orientation-Dependent Mobility Engineering

Besides substrate-induced and process-induced stress engineering, wafer substrate orientation and channel orientation can improve mobility. Different surface orientation and direction of applied field for different in-plane stress provide different interaction with carrier transport [29]. Another approach to improve on the mobility and current drive capabilities of Si involves the exploitation of the mobility anisotropy of Si. Fig. 1.5 shows the mobility of electrons and holes for different surface orientations and in the case of the (110) surface, two different in-plane channel directions. It is well-known that electron mobility is highest for the traditional Si substrate/transistor configu-

ration, which includes a (100) surface with a <110> channel direction, while hole mobility is highest for a (110) surface with a <110> channel direction. A hybrid substrate configuration [28] has been proposed to allow for the ultimate mobility configuration for each carrier. It should also be noted that even without a change in surface orientation from the traditional (100) surface, hole mobility has been reported to increase via rotation of the channel direction by $45^o$ from the <110> direction to the <100> direction [30, 31, 32]. This effect has been reported to be used in 90-nm node products [33]. As with conventional Si surface and channel geometries, the introduction of strain in tandem with different substrate/channel orientations is under investigation [34, 35].

## 1.3 Gate Dielectrics on Engineered Substrates

Perhaps the most significant sign of the change in the Si industry is that the new materials, such as high-k materials, are being taken seriously as replacements for $SiO_2$ as a gate dielectric and isolation layer. The SiO(N) gate insulator in Si-based MOSFETs has been scaled close to its physical thickness limit. To further increase chip performance by using novel materials, the focus has been on high-permittivity (high-k) gate insulators and on metal gate electrodes; increasing the carrier mobility in the channel is the next challenge.

New materials such as high permittivity dielectrics will eventually replace the thermal gate oxide or oxynitride to reduce gate leakage and further improve the drive current with scalable equivalent oxide thickness (EOT). A good high-k material will need to be thermally and chemically stable, have reasonable dielectric value, and result in good silicon and gate electrode interface properties. The high-k dielectric thickness needs to be carefully chosen to avoid interfering with the scalability of the device.

The concurrent reduction of the gate oxide thickness with the channel length results in a large tunneling current. As the gate oxide scales below 15 Å, the gate leakage current is becoming larger than the off-state current of the device required for low power consumption at the 65-nm technology node and beyond. Excessive gate leakage current due to ultra-thin oxide films will necessitate the introduction of high-k material. The technology barriers with high-k dielectric and metal gate are:

- Scaling $SiO_2$ gate dielectrics
- Doping limits in silicon and polysilicon
- Ion implantation limits
- Threshold voltage scaling limits
- Bulk CMOS structure limits

## 1.4 Strained-Si Technology: Process Integration

The next consideration for strained-Si is the fabrication of strained-Si devices and circuits. Strained-Si technology provides great performance benefits for CMOS, but as with any new technology that is introduced to a manufacturing infrastructure, the ease with which it is adopted is very dependent on the number and severity of changes the new technology requires. Strained-Si technology has attracted the attention of chip designers since the mid-1990s. The strained-Si is a technology based on growing a layer of SiGe alloy on top of a Si wafer. About 4.18% lattice mismatch between Si and Ge causes crystalline defects in SiGe film made by epitaxial deposition. During the deposition of SiGe layers, it is important to avoid a surface defect problem called crosshatch. The ability to neutralize the crosshatch problem enables one to deposit up to 100% Ge in the SiGe layer. Amberwave's strained-Si technology eliminates the surface roughness and defect problems that plague other strained-Si technologies.

Although the majority of early studies have mainly focused on the mobility and performance of long-channel devices, recent studies have shown possible performance improvement of strained-Si MOSFETs with even sub-100 nm gate lengths. The typical process steps used for the fabrication of CMOS circuits were used for strained-Si substrates, which means that no equipment changes or specialty processes are required. The semiconductor R&D community has striven for more than 10 years to develop strained-Si as a cost-effective manufacturing alternative.

One feature of strained-Si technology that distinguishes it most from bulk-Si technology is the fact that the thickness of strained-Si layer, which is limited in thickness due to critical thickness constraints, cannot be stripped indiscriminately. This limits the use of aggressive cleaning strips and other processes that are commonplace in the Si industry. In practice, as long as approximately 5-nm of strained-Si remains at the end of the processing sequence, the transistors will exhibit optimal performance.

It is important to note that the equipment and processes used in standard epitaxial Si wafer processing can be readily implemented for the manufacture of strained-Si substrates and no new manufacturing infrastructure is required. However, there may be specific modules that need to be modified, as the material properties of the strained-Si substrate are somewhat different from bulk-Si. For example, dopant activation and dopant diffusion characteristics are altered due to the presence of Ge in the substrate. This results in different thermal budgets for strained-Si technology.

The presence of the SiGe layer in strained-Si substrate leads to several material- and integration-related challenges. For example, SiGe strain relaxation layer induces high density of defects in strained-Si. Dopant diffusion rate in SiGe is significantly different from that in Si. Boron diffusion is retarded

while arsenic diffusion is enhanced. Additional efforts in junction engineering to control short channel effects and to set the MOSFET threshold voltages to the desired values are required. Significant device self-heating is also observed in strained-Si/SiGe devices because of the lower thermal conductivity in SiGe. The strained-Si technology faces three key challenges: short-channel effect, silicon dislocations, and keeping the strain level during manufacturing. However, strained-Si technology has not shown any known yield-limiting issues so far. Although early research has focused primarily on n-MOSFET devices for the ease of electron mobility, however, recent demonstrations in current improvements are of more than 40% in p-MOSFETs.

Fabrication of strained-Si circuits with symmetric n- and p-MOSFET with high drive current has been demonstrated. It is also possible to build dual-gate devices by combining strained-Si with an SOI substrate in an attempt to overcome difficult performance problems with CMOS scaling. As the infrastructure for circuit design does not change with strained-Si technology, one can expect the SPICE models and design libraries similar to those available today from the fabrication will remain almost the same.

## 1.5 Nonclassical CMOS Structures

So far, industry focus has concentrated on transistor scaling via shortening of the transistor gate length and transistors fabricated at technology nodes below 0.13 $\mu$m show a rapid decrease in drive current. Existing capabilities of conventional Si substrates will not be able to meet ITRS roadmap specifications. Fig. 1.6 is a plot of device delay vs. the power consumed for various material systems [36]. Fig. 1.6 shows a plot of I/CV, essentially the figure of merit for digital ICs, which is directly correlated to power-delay product. It is apparent that unless drive current enhancements are introduced, productive scaling would end [37]. New substrate materials are being implemented to continue this scaling [38, 39, 40, 41]. With the use of biaxial strained-Si layers, drive current enhancements were shown to be achievable with implementation of strain, as seen in the strained-Si curve for I/CV, based on the assumption of a gate-independent enhancement of 20%. For the present, it is thus possible to stay at a given technology and obtain increase in drive current without further scaling.

For over four decades now, the semiconductor industry has relied mostly on shrinking transistor geometries for improvements in both circuit performance and density, resulting in a consistent reduction of cost per function for each new technology node. These include strained-Si, SiGe, high-k gate dielectrics, metal gates, ultrathin body fully depleted SOI (FDSOI), and then multi-gated devices (e.g., FinFET). As CMOS technology continues to scale down, the

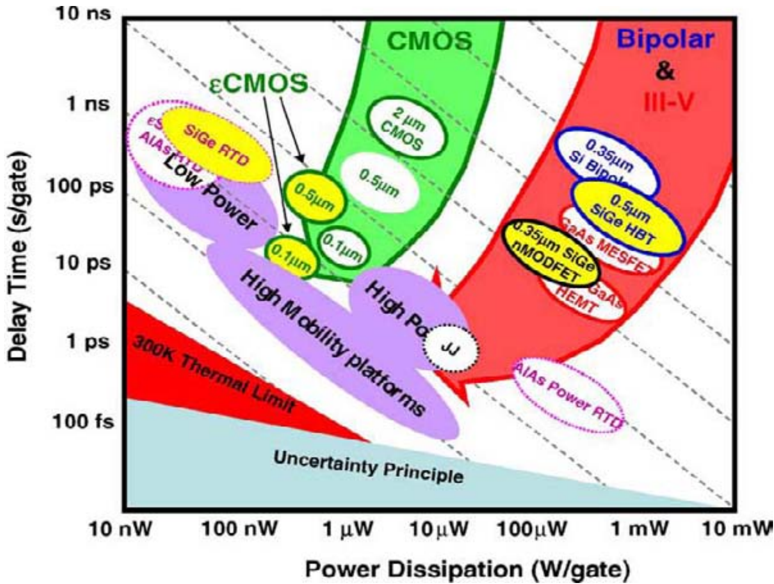



**FIGURE 1.6**
Power-delay product for digital integrated circuits. This diagram shows how both materials and device affect the power-delay product. The dashed lines are lines of equal switching energy. Regions are indicated that can benefit most from engineered substrates with higher mobility layers (i.e., materials with lower power-delay product). Low power, high power, and scaled high mobility platform regions are indicated. After D. J. Paul, *Thin Solid Films*, Vol. 321, 1998(172-180). With permission.

FDSOI and double-gate FET (DGFET) technologies offer potential solutions to control short-channel effects by reducing the silicon film thickness and a concurrent scaling of the buried oxide thickness.

Nonclassical MOS transistor structures such as thin-body silicon-on-insulator and double-gate transistors can be scaled more aggressively than the classical bulk-Si structure, and hence may be adapted within the next five years, at the 65-nm technology node (25-nm physical gate length). These advanced structures have distinctly different materials and process technology requirements and associated challenges.

To continue MOSFET scaling to less than $L_g$ = 20 nm, it is quite likely that the device structure will change to advanced nonclassical CMOS such as multiple-gate, ultrathin body (UTB) MOSFETs. In these devices, various "technology boosters" such as mobility enhancement by strained-Si, elevated source/drain, high-k gate dielectric, and metal gate electrode, will likely be simultaneously implemented with the new device structure. In UTB MOSFETs having less than 10-nm Si thickness, various quantum effects will impact the electric characteristics. Toward the end of the roadmap, devices will increas-

ingly be operated in the quasiballistic mode, where the current gain will be enhanced by parameters different from those currently known. Choice of the optimum device structures, their physical characterization, and construction of cost-effective processing flows will become very important along with construction of their circuit architecture. Nanotubes and nanowires may finally be long-term alternatives due to their aspect ratios, which bring them close to being ideal one-dimensional (1D) systems.

However, each of the aforementioned components can result in incremental improvements to CMOS devices; it is anticipated that once the device shrinks beyond the 90-nm node, multiple combinations of these components will be required to achieve the requisite performance improvement and power consumption reduction.

## 1.6 Strain-Engineered Hetero-FETs: Modeling and Simulation

With the adoption of stress engineering as a key component of the new device architectures, modeling of the front-end stress and associated stress enhancement of mobility has become increasingly important for scaling of CMOS devices [14, 42]. The sources of stress in silicon technology come from a variety of strain generated by different processes such as oxidation, thermal mismatch, intrinsic stress, and dopants. Stress modeling is becoming a very important part of the simulation methodology needed to help evaluate and engineer improved device performance due to stress. However, it must be used along with calibrated process simulation for doping profiles and a device simulator that correctly models the change in device characteristics due to stress.

In the area of modeling, as atomic distances are approached, understanding and modeling the complex operation of nano transistors become extremely challenging. Bottoms-up chemical perspective of engineering molecular architecture and correlating precise molecular components with device function will offer new insights into fundamental semiconductor dynamics, facilitate cost-effective manufacturing, and stimulate new possibilities. In addition, a new concern arises about whether the scaling could continue. Also aggressive scaling of the gate length for high-performance applications makes device parameter optimization quite difficult. Ultrashallow junction formation for suppressing the short-channel effect cannot be achieved without incurring a significant increase of parasitic resistance.

Emerging nanoscale semiconductor physical phenomena need to include the following:

- Microscale to nanoscale dimensions

- Single dopant effects
- Surface/interface dominance
- Interface-to-interface interactions in nanoscale devices
- Materials fabrication - reaction kinetics and defects and dislocation
- Nonperiodicity effects
- High field effects
- Single-particle transport phenomena

Advanced MOSFET device physics should include:

- MOSFET structure design
- Short-channel effects
- Electric field distributions in channel and source/drain
- Equilibrium bulk and interface properties which determine performance
- Gate oxide/semiconductor interface control and engineering
- Transient charge carrier transport

## 1.7 Summary

Attempts are being made to identify the avenues of research solutions proposed by ITRS to address the future technology needs extending front-end fabrication of CMOS from the 65-nm node to its ultimate limit at or beyond the 22-nm node. Emphasis is being placed on front-end processing and materials, and device structures, along with the supporting characterization, modeling, and simulation techniques. A multidisciplinary effort is required to lead this exciting field to fruition.

In this chapter, we have presented several issues involved with new materials and structures being introduced into Si CMOS technology, and presented some possible potential solutions to enable high-performance and low-power CMOS for the 45-nm technology node and beyond. Various technology innovations ("technology boosters"), such as strained-Si, metal gate, ultrathin body SOI MOSFETs, and multiple-gate MOSFETs, including FinFETs, have been introduced. The proposed solutions should be pursued concurrently with circuit design and architecture improvements, particularly to manage power dissipation.

In the view of the limitations to increase device performance through continued scaling, manufacturers now may find an opportunity in strained-Si to get a reasonable performance enhancement with a relatively simple change in starting materials and hence the cost. However, there are many reasons to believe that continued device scaling will not be as straightforward in the future as it has been in the past. Practical and/or fundamental limits are being approached and substantial changes to device technologies and structures are

going to be required. While "inventions" and new materials have largely not been needed for the past 30 years, they surely will be needed over the next 15 years. This period will likely be the most challenging that the semiconductor industry has faced because it is likely that during this period we will really understand how far Moore's law can be extended.

With highly creative discoveries and developments of new materials, processes, and transistor structures, many of the challenges have been met so far. Communication between scientists and engineers working in the fields of (a) physics, chemistry, growth, and fabrication of novel substrate materials, (b) device fabrication, (c) characterization, and (d) modeling needs to be increased. However, it may be noted that most of the industry has not yet embraced most of the advanced substrates and device structures that have been demonstrated at the research level.

## References


[1] J. E. Lilienfeld, "Method and apparatus for controlling electric currents," US Patent 1745175, 1925.

[2] J. E. Lilienfeld, US Patent 1900018, 1928.

[3] O. Heil, British Patent 439457, 1935.

[4] D. Kahng, US Patent 3102230, 1960.

[5] J. D. Plummer, M. D. Deal, and P. B. Griffin, *Silicon VLSI Technology - Fundamentals, Practice and Modeling.* Prentice Hall, Upper Saddle River, NJ, 2000.

[6] P. Bai, C. Auth, S. Balakrishnan, M. Bost, R. Brain, V. Chikarmane, R. Heussner, M. Hussein, J. Hwang, D. Ingerly, R. James, J. Jeong, C. Kenyon, E. Lee, S-H. Lee, N. Lindert, M. Liu, Z. Ma, T. Marieb, A. Murthy, R. Nagisetty, S. Natarajan, J. Neirynck, A. Ott, C. Parker, J. Sebastian, R. Shaheed, S. Sivakumar, J. Steigerwald, S. Tyagi, C. Weber, B. Woolery, A. Yeoh, K. Zhang, and M. Bohr, "A 65 nm logic technology featuring 35 nm gate lengths, enhanced channel strain, 8 Cu interconnect layers, low-k ILD and 0.57 $\mu m^2$ SRAM Cell," *IEEE IEDM Tech. Dig.*, pp. 657–660, 2004.

[7] S. E. Thompson, R. S. Chau, T. Ghani, K. Mistry, S. Tyagi, and M. T. Bohr, "In search of 'forever' continued transistor scaling one new material at a time," *IEEE Trans. Semicond. Manufacturing*, vol. 18, pp. 26–36, 2005.

[8] T. P. Pearsall and J. C. Bean, "Enhancement-mode and depletion-mode p-channel $Ge_xSi_{1-x}$ modulation-doped FETs," *IEEE Electron Dev. Lett.*, vol. 7, pp. 308–310, 1986.

[9] C. K. Maiti, L. K. Bera, and S. Chattopadhyay, "Strained-Si heterostructure field effect transistors," *Semicond. Sci. Technol.*, vol. 13, pp. 1225–1246, 1998.

[10] M. Houssa, *High-k Gate dielectrics.* Institute of Physics Publishing, Bristol, 2003.

[11] A. V. Y. Thean, T. White, M. Sadaka, L. McCormick, M. Ramon, R. Mora, P. Beckage, M. Canonico, X.-D. Wang, S. Zollner, S. Murphy, V. van Der Pas, M. Zavala, R. Noble, O. Zia, L.-G. Kang, V. Kolagunta, N. Cave, J. Cheek, M. Mendicino, B.-Y. Nguyen, M. Orlowski, S. Venkatesan, J. Mogab, C. H. Chang, Y. H. Chiu, H. C. Tuan, Y. C. See, M. S. Liang, Y. C. Sun, I. Cayrefourcq, F. Metral, M. Kennard, and C. Mazure, "Performance of super-critical strained-Si directly on insulator (SC-SSOI) CMOS based on high-performance PD-SOI technology," *IEEE VLSI Tech. Symp. Dig.*, pp. 134–135, 2005.

[12] K. Rim, S. Koester, M. Hargrove, J. Chu, P. M. Mooney, J. Ott, T. Kanarsky, P. Ronsheim, M. Ieong, A. Grill, and H.-S. P. Wong, "Strained Si NMOSFETs for high performance CMOS technology," *VLSI Technology*, pp. 59–60, 2001.

[13] H. S. Yang, R. Malik, S. Narasimha, Y. Li, R. Divakaruni, P. Agnello, S. Allen, A. Antreasyan, J. C. Arnold, K. Bandy, M. Belyansky, A. Bonnoit, G. Bronner, V. Chan, X. Chen, Z. Chen, D. Chidambarrao, A. Chou, W. Clark, S. W. Crowder, B. Engel, H. Harifuchi, S. F. Huang, R. Jagannathan, F. F. Jamin, Y. Kohyama, H. Kuroda, C. W. Lai, H. K. Lee, W.-H. Lee, E. H. Lim, W. Lai, A. Mallikarjunan, K. Matsumoto, A. McKnight, J. Nayak, H. Y. Ng, S. Panda, R. Rengarajan, M. Steigerwalt, S. Subbanna, K. Subramanian, J. Sudijono, G. Sudo, S.-P. Sun, B. Tessier, Y. Toyoshima, P. Tran, R. Wise, R. Wong, I. Y. Yang, C. H. Wann, L. T. Su, M. Horstmann, Th. Feudel, A. Wei, K. Frohberg, G. Burbach, M. Gerhardt, M. Lenski, R. Stephan, K. Wieczorek, M. Schaller, H. Salz, J. Hohage, H. Ruelke, J. Klais, P. Huebler, S. Luning, R. van Bentum, G. Grasshoff, C. Schwan, E. Ehrichs, S. Goad, J. Buller, S. Krishnan, D. Greenlaw, M. Raab, and N. Kepler, "Dual stress liner for high performance sub-45 nm gate length SOI CMOS manufacturing," *IEEE IEDM Tech. Dig.*, pp. 1075–1077, 2004.

[14] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson, and M. Bohr, "A 90 nm high volume manu-

facturing logic technology featuring novel 45 nm gate length strained silicon CMOS," *IEEE IEDM Tech. Dig.*, pp. 978–980, 2003.

[15] S. Datta, G. Dewey, M. Doczy, B. S. Doyle, B. Jin, J. Kavalieros, R. Kotlyar, M. Metz, N. Zelick, and R. Chau, "High mobility Si/SiGe strained channel MOS transistors with $HfO_2$/TiN gate stack," *IEEE IEDM Tech. Dig.*, pp. 653–656, 2003.

[16] T. Sanuki, A. Oishi, Y. Morimasa, S. Aota, T. Kinoshita, R. Hasumi, Y. Takegawa, K. Isobe, H. Yoshimura, M. Iwai, K. Sunouchi, and T. Noguchi, "Scalability of strained silicon CMOSFET and high drive current enhancement in the 40 nm gate length technology," *IEEE IEDM Tech. Dig.*, pp. 65–68, 2003.

[17] M. Horstmann, D. Greenlaw, Th. Feudel, A. Wei, K. Frohberg, G. Burbach, M. Gerhardt, M. Lenski, R. Stephan, K. Wieczorek, M. Schaller, J. Hohage, H. Ruelke, J. Klais, P. Huebler, S. Luning, R. van Bentum, G. Grasshoff, C. Schwan, J. Cheek, J. Buller, S. Krishnan, M. Raab, and N. Kepler, "Sub-50 nm gate length SOI transistor development for high performance microprocessors," *Mat. Sci. and Eng. B*, vol. 114-115, pp. 3–8, 2004.

[18] J. Goo, Q. Xiang, Y. Takamura, H. Wang, J. Pan, F. Arasnia, E. N. Paton, P. Besser, M. V. Sidorov, E. Adem, A. Lochtefeld, G. Braithwaite, M. T. Currie, R. Hammond, M. T. Bulsara, and M. Lin, "Scalability of strained-Si nMOSFETs down to 25 nm gate length," *IEEE Electron Dev. Lett.*, vol. 24, pp. 351–353, 2003.

[19] H. C.-H. Wang, Y.-P. Wang, S.-J. Chen, C.-H. Ge, S. M. Ting, J.-Y. Kung, R.-L. Hwang, H.-K. Chiu, L. C. Sheu, P.-Y. Tsai, L.-G. Yao, S.-C. Chen, H.-J. Tao, Y.-C. Yeo, W.-C. Lee, and C. Hu, "Substrate-strained silicon technology: process integration," *IEEE IEDM Tech. Dig.*, pp. 61–64, 2003.

[20] J. W. Matthews and A. E. Blakeslee, "Defects in epitaxial multilayers - I. Misfit dislocations," *J. Crystal Growth*, vol. 27, pp. 118–125, 1974.

[21] T. Sato, Y. Takeishi, H. Hara, and Y. Okamoto, "Mobility anisotropy of electrons in inversion layers on oxidized silicon surfaces," *Phys. Rev. B*, vol. 4, pp. 1950–1960, 1971.

[22] S. Takagi, T. Mizuno, N. Sugiyama, T. Tezuka, and A. Kurobe, "Strained-Si-on-insulator (strained-SOI) MOSFETs - concept, structures and device characteristics," *IEICE Trans. on Electron.*, vol. E84-C, pp. 1043–1050, 2000.

[23] H.-S. P. Wong, "Beyond the conventional transistor," *IBM J. Res. & Dev.*, vol. 46, pp. 133–168, 2002.

[24] T. A. Langdo, A. Lochtefeld, M. T. Currie, R. Hammond, V. K. Yang, J. A. Carlin, C. J. Vineis, G. Braithwaite, H. Badawi, M. T. Bulsara, and E. A. Fitzgerald, "Preparation of novel SiGe-free strained-Si on insulator substrates," *Proc. IEEE Int. SOI Conf.*, pp. 211–212, 2002.

[25] K. Rim, R. Anderson, D. Boyd, F. Cardone, K. Chan, H. Chen, S. Christansen, J. Chu, K. Jenkins, T. Kanarsky, S. Koester, B. H. Lee, K. Lee, V. Mazzeo, A. Mocuta, D. Mocuta, P. M. Mooney, P. Oldiges, J. Ott, P. Ronsheim, R. Roy, A. Steegen, M. Yang, H. Zhu, M. Ieong, and H.-S. P. Wong, "Strained-Si CMOS (SS CMOS) technology: opportunities and challenges," *Solid-State Electron.*, vol. 47, pp. 1133–1139, 2003.

[26] K. Rim, K. Chan, L. Shi, D. Boyd, J. Ott, N. Klymko, F. Cardone, L. Tai, S. Koester, M. Cobb, D. Canaperi, B. To, E. Duch, I. Babich, R. Carruthers, P. Saunders, G. Walker, Y. Zhang, M. Steen, and M. Ieon, "Fabrication and mobility characteristics of ultrathin strained-Si directly on insulator (SSDOI) MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 49–52, 2003.

[27] K. Rim, J. Chu, H. Chen, K. A. Jenkins, T. Kanarsky, K. Lee, A. Mocuta, H. Zhu, R. Roy, J. Newbury, J. Ott, K. Petrarca, P. Mooney, D. Lacey, S. Koester, K. Chan, D. Boyd, M. Leong, and H.-S. Wong, "Characteristics and device design of sub-100 nm strained-Si n- and p-MOSFETs," *IEEE VLSI Tech. Symp. Dig.*, pp. 98–99, 2002.

[28] M. Yang, M. Ieong, L. Shi, K. Chan, V. Chan, A. Chou, E. Gusev, K. Jenkins, D. Boyd, Y. Ninomiya, D. Pendleton, Y. Surpris, D. Heenan, J. Ott, K. Guarini, C. D'Emic, M. Cobb, P. Mooney, B. To, N. Rovedo, J. Benedict, R. Mo, and H. Ng, "High performance CMOS fabricated on hybrid substrate with different crystal orientations," *IEEE IEDM Tech. Dig.*, pp. 453–456, 2003.

[29] M. V. Fischetti, Z. Ren, P. M. Solomon, M. Yang, and K. Rim, "Six-band k·p calculation of the hole mobility in silicon inversion layers: Dependence on surface orientation, strain, and silicon thickness," *J. Appl. Phys.*, vol. 94, pp. 1079–1095, 2003.

[30] H. Sayama, Y. Nishida, H. Oda, T. Oishi, S. Shimizu, T. Kunikiyo, K. Sonoda, Y. Inoue, and M. Inuishi, "Effect of <100> channel direction for high performance SCE immune pMOSFET with less than 0.15 $\mu$m gate length," *IEEE IEDM Tech. Dig.*, pp. 657–660, 1999.

[31] T. Matsumoto, S. Maeda, H. Dang, T. Uchida, K. Ota, Y. Hirano, H. Sayama, T. Iwamatsu, T. Ipposhi, H. Oda, S. Maegawa, Y. Inoue, and T. Nishimura, "Novel SOI wafer engineering using low stress and high mobility CMOSFET with <100>-channel for embedded RF/Analog applications," *IEEE IEDM Tech. Dig.*, pp. 663–666, 2002.

[32] T. Komoda, A. Oishi, T. Sanuki, K. Kasai, H. Yoshimura, K. Ohno, M. Iwai, M. Saito, F. Matsuoka, N. Nagashima, and T. Noguchi, "Mobility improvement for 45 nm node by combination of optimized stress control and channel orientation design," *IEEE IEDM Tech. Dig.*, pp. 217–220, 2004.

[33] D. James, "2004 - The year of 90 nm: a review of 90 nm devices," *IEEE/SEMI ASMC Proc.*, pp. 1–6, 2005.

[34] T. Mizuno, N. Sugiyama, T. Tezuka, and S. Takagi, "[110] strained-SOI n-MOSFETs with higher electron mobility," *IEEE Electron Dev. Lett.*, vol. 24, pp. 266–268, 2003.

[35] T. Mizuno, N. Sugiyama, T. Tezuka, Y. Moriyama, S. Nakaharai, T. Maeda, and S. Takagi, "Physical mechanism for high hole mobility in (110)-surface strained- and unstrained-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 809–812, 2003.

[36] D. J. Paul, "Silicon germanium heterostructures in electronics: the present and the future," *Thin Solid Films*, vol. 321, pp. 172–180, 1998.

[37] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors." San Jose, Calif., 2005.

[38] D. K. Nayak, K. Goto, A. Yutani, J. Murota, and Y. Shiraki, "High-mobility strained-Si PMOSFETs," *IEEE Trans. Electron Dev.*, vol. 43, pp. 1709–1715, 1996.

[39] D. K. Nayak, J. C. S. Woo, J. S. Park, K. L. Wang, and K. P. MacWilliams, "High-mobility p-channel metal-oxide-semiconductor field-effect transistor on strained-Si," *Appl. Phys. Lett.*, vol. 62, pp. 2853–2855, 1993.

[40] C. K. Maiti, L. K. Bera, S. S. Dey, D. K. Nayak, and N. B. Chakrabarti, "Hole mobility enhancement in strained-Si p-MOSFETs under high vertical fields," *Solid-State Electron.*, vol. 41, pp. 1863–1869, 1997.

[41] J. Welser, J. L. Hoyt, S. Takagi, and J. F. Gibbons, "Strain dependence of the performance enhancement in strained-Si n-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 373–376, 1994.

[42] C.-H. Ge, C.-C. Lin, C.-H. Ko, C.-C. Huang, Y.-C. Huang, B.-W. Chan, B.-C. Perng, C.-C. Sheu, P.-Y. Tsai, L.-G. Yao, C.-L. Wu, T.-L. Lee, C.-J. Chen, C.-T. Wang, S.-C. Lin, Y.-C. Yeo, and C. Hu, "Process-strained Si (PSS) CMOS technology featuring 3D strain engineering," *IEEE IEDM Tech. Dig.*, pp. 73–76, 2003.

# 2

# *Strain Engineering in Microelectronics*

After more than three decades of continued progress in CMOS devices technology, especially aggressive scaling in the last few years, the CMOS scaling is now approaching the fundamental limits. As we enter the nanometer regime, stress from standard process steps such as source/drain doping introduce significant stress in the channel of MOSFETs. With the continuing reduction of device dimensions, the impact of process-induced stress on device performance is becoming increasingly important. Besides scaling, several innovative mobility enhancement techniques are being attempted to maintain the CMOS performance improvement. Mobility enhancement is attractive because it improves device performance without device scaling. However, continued miniaturization increases device complexity and internal mechanical stress. The stress may cause dislocations, film cracking, and degradation of gate oxide quality. To reduce or avoid the buildup of stress during processing, it is important to identify the stress causing process steps, and locate high-stress areas with high spatial resolution.

It has been shown that the introduction of a compressive and/or tensile stress in the Si channel can improve the mobility of both carrier types and, in turn, increase the n- and p-MOSFET drive currents [1, 2]. Various methods have also been proposed to induce the desired biaxial tensile strain in the channel such as via the use of a SiGe virtual substrate [3], tensile films [4], or mechanical force [5, 6].

Thin films such as silicon nitride, silicon dioxide, and polysilicon are frequently encountered in device fabrication; these films contain intrinsic stresses as a result of the deposition process. With the continuing reduction of device dimensions, the impact of the isolation process on device performance also becomes increasingly important. A major concern when scaling LOCOS and shallow trench isolation (STI) structures is the buildup of localized stress near the isolation edge, often leading to dislocation formation and degrading device characteristics.

Compressive stresses are introduced by shallow trench isolation (STI) toward channels longitudinally and laterally. At STI, a large volume of oxide is deposited in a trench to create isolation structures. Residual stress from STI can arise from multiple factors upon cooling from high temperatures above the oxide glass transition. The volume contraction between the silicon and oxide results in stress from thermal expansion. Further, densification of the oxide at high temperatures due to changes in bonding can build up intrinsic residual

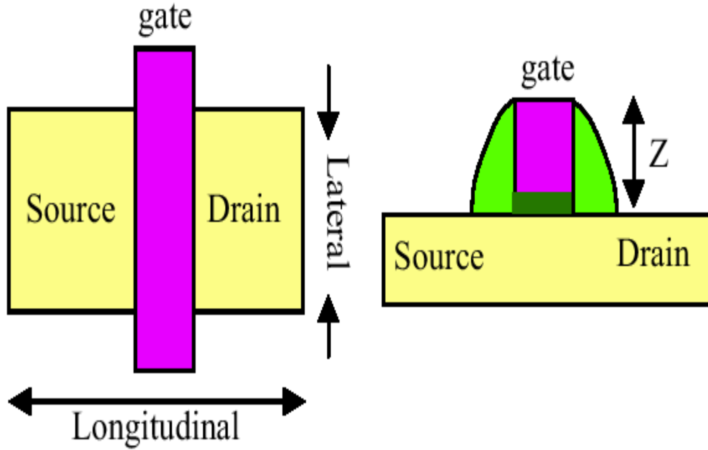



**FIGURE 2.1**
Effects of stress on CMOS performance. After V. Chan et al., *IEEE 2005 CICC Conf.*, 2005(667-674). With permission.

stress in the STI oxides. This residual stress is compressive for most commonly used oxide filling materials. Compressive stresses can be introduced by shallow trench isolation toward channels longitudinally and laterally. n- and p-MOSFETs show different enhancements in three directions, namely longitudinal, lateral and Si depth directions (see Table 2.1) for applied channel strain as shown in Fig. 2.1 which shows the possible directions of uniaxial stresses and their effects on n- and p-MOSFETs. Mechanical stress breaks crystal symmetry and removes the two-fold and six-fold degeneracy of the valence and conduction bands, respectively. The strains created at the isolation edge decay monotonically toward the middle of the channel, and the distance between the gate edge and isolation edge determines the actual strain in the channel.

The introduction of strain changes the mechanical, electrical (band structure and mobility), and chemical (diffusion and activation) properties of a semiconductor. The various effects of stress and strain on silicon and silicon technology have been studied since the 1950s. Most significant to silicon technology are the changes in band gap, effective mass, mobility, diffusivity of dopants, and oxidation rates [7]. The strain effects on mobility were found to be anisotropic and carrier effects were found to be different for bulk silicon and inversion layers [8, 9].

Different mechanical stress can be introduced through the substrate and/or during the fabrication process, which are discussed below. The influence of lattice defects and externally applied stress on the electrical performance of integrated circuits has been an active area of research for many years. In the

**TABLE 2.1**
Effects of stress on CMOS performance. After V. Chan et al., *IEEE 2005 CICC Conf.*, 2005(667-674). With permission.

| | nMOS | pMOS |
|---|---|---|
| Longitudinal X | Tensile | Compressive |
| Lateral Y | Tensile | Tensile |
| Si Depth Z | Compressive | Tensile |

past, a major concern has been the buildup of localized stress near the isolation edge, often leading to dislocation formation and degradation in device characteristics when scaling the LOCOS and shallow trench isolation structures are introduced. Changes in electron and hole mobility due to stress from local oxidation of silicon (LOCOS) and shallow trench isolation were reported [10, 11]. However, it was not until deep submicrometer technologies were developed that these effects were observed experimentally.

Since the early 1990s, strained-Si on silicon germanium (SiGe) substrates are being explored in an effort to boost MOSFET performance [3, 12, 13]. The mobility advantage that strain offers at no significant additional processing cost makes it the most important candidate among various choices for performance-enhanced devices. Approaches to introduce device strain have either been global, or wafer-level, where the strain exists within the starting substrate [14, 15, 16], or local where strain is introduced during transistor fabrication using local stressors. Now, localized strain is being deliberately introduced in a variety of ways to enhance carrier mobility, although the strain from localized source decays rapidly away from the stressor. Today, almost every semiconductor manufacturer has announced their version of strained CMOS [4, 17, 18, 19, 20, 21, 22].

Strain is characterized by its two major varieties: uniaxial strain (one surface dimension is stressed) vs. biaxial strain (both surface dimensions are stressed). SiGe-based biaxial strain has been the most prevalent wafer-level strained substrate but has encountered integration issues due to its high defect levels. Generally, global and local strain engineering techniques are applicable to both bulk-Si and silicon-on-insulator device architectures: two important trends that can be observed in the data in Table 1.2 and the literature at large. Local strain engineering techniques have been demonstrated concurrently for n- and p-MOSFET integration schemes, while global strain engineering techniques have generally been optimized for n-MOSFETs only. n-MOSFET enhancements with global strained-Si technology are typically twice that of local strain engineering techniques. Significant p-MOSFET enhancements have so far been based upon local strain engineering techniques.

As silicon is strained and subjected to high temperature thermal processing, it undergoes mechanical deformations that create defects, which may significantly limit yield. The engineering challenge is to understand and op-

timize how the stress/strain originating from the various techniques interact with one another. Engineers have to manipulate these properties of silicon to balance the performance gains against defect generation. From process integration point of view, SiGe materials offer several new challenges to understanding. In the nanometer regime, biaxial stress has been the conventional method to strain the MOSFET channel. Recent studies in the field of uniaxial process-induced strain have revealed significant advantages over its biaxial counterpart.

In this chapter, we shall introduce the role of strain and stress engineering in microelectronics. We shall elucidate the current understanding and ongoing research efforts in strain engineering. The main focus of this chapter is to provide a survey on the applications of stress and strain in silicon technology. We shall focus on strained-Si and discuss various methods of applying stress to enhance device performance. Stress-transfer techniques (such as etch-stop liner and strained-SiGe) using extra integration process into an existing baseline process and mechanically induced channel stress from shallow trench isolation structures will be discussed.

## 2.1 Stress Induced During Manufacturing

Semiconductor device performance has been dramatically improved so far by scaling it to nanometer dimensions. In fabrication, multiple material layers with differing thermal coefficients of expansion and hence the stress levels developed in close proximity of the devices and in the channel region have become a cause of serious concern. Among the semiconductors, although silicon has rather poor transport properties (mobilities and diffusion coefficients for electron and holes), its success as channel material is due to the excellent interface property of the Si-$SiO_2$ system. In order to sustain Moore's law for over the last four decades, many material and design aspects of the conventional MOSFETs have been attempted and lost, except the use of Si-$SiO_2$ material system as channel material and gate insulator. Materials commonly used in microelectronics have very different elastic properties. Mostly they obey complex rheological laws (viscoelastic or elastoplastic), and it is important to know the temperature dependence of some rheological characteristics like viscosity or yield stress.

In the range of temperatures used in the front-end processes (700-1100$^o$C), and in the range of stresses measured in thin film materials (20 MPa-2 GPa), microelectronics materials show very different mechanical behaviors, from elastic (e.g., Si) to viscous (e.g., highly stressed $SiO_2$ at high temperature). Furthermore, at these temperatures and/or stress levels, the amorphous materials may exhibit viscoelastic behavior and the poly-crystalline materials may

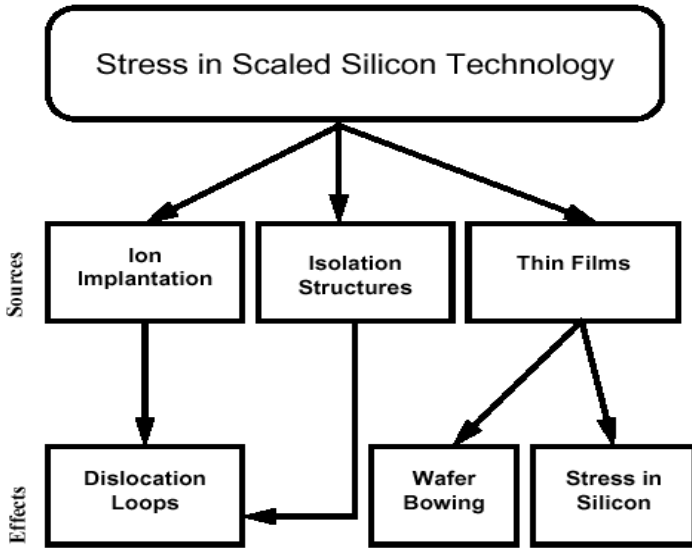



**FIGURE 2.2**
Sources of various stresses in silicon microelectronics. After S. Chaudhry, "Analysis and Modeling of Stress Related Effects in Scaled Silicon Technology," Doctoral Dissertation, University of Florida, 1996.

have elastoplastic behavior as well. Residual stresses are critical in Si technology and can be experimentally determined using either optical techniques allowing the calculation of the substrate curvature, or directly by micro-Raman spectroscopy or x-ray/electron diffraction techniques. Fig. 2.2 is a summary of various stresses encountered in silicon microelectronics. The stresses can generally be classified into five groups:

(a) the temperature change and the thermal growth of material (i.e., silicon dioxide and silicides),

(b) the etching of materials,

(c) the deposition processes which create intrinsic stresses,

(d) the structural changes of deposited films due to densification, hydration or dehydration, crystallization, and

(e) the implantation of dopants. The total measured stress is the sum of these five components.

Thermal stresses are induced by the mismatch of the thermal expansion coefficients between different materials. They arise when the system undergoes a temperature variation during the various stages of the integrated circuit processing. By introducing a number of assumptions; (i) materials are isotropic and homogeneous and (ii) there is no temperature gradient in the film or in the substrate, the thermal strain, $\epsilon$ induced by a variation in temperature can be calculated from

$$\epsilon = \alpha(T) \cdot \Delta T \tag{2.1}$$

where $\alpha$(T) is the thermal expansion coefficient of the film. For silicon, it has been observed that this coefficient is temperature dependent. However, for most materials used in silicon IC processes the dependence in temperature is either negligible or unknown.

Thermal oxidation of nonplanar silicon surfaces is frequently used in silicon technology. During the oxidation process, the volume expansion, resulting from silicon to $SiO_2$ conversion, is accommodated by the viscoelastic flow of oxide, while the silicon substrate may undergo elastoplastic deformations, sometimes with crystal dislocation generation. Oxidation-induced stress analyses have been extensively proposed in the literature during the past few years. Several approaches have been reported in order to accurately reproduce the oxidation mechanism and to calculate its corresponding stress field which takes into account the lateral growth of oxide, in agreement with experimental evidence showing wafer bending during planar oxidations.

In MOSFET fabrication, high dose ion implantation is generally used to dope source/drain regions to desired junction depths. The implanted region becomes amorphized and upon subsequent annealing dislocations form at the original amorphous crystalline interface. When a dopant atom is introduced in silicon substrate through ion implantation or diffusion, a local lattice expansion or contraction will occur depending on the varying atomic sizes and bond lengths of the atoms. The mismatch between the covalent radii of silicon and dopant atoms results in localized lattice deformation of the substrate. Following Vegard's law, one can extract the relative lattice variation ($\Delta$a/a) induced for a given concentration ($C_d$) of dopants

$$\Delta a/a = \beta \cdot C_d \tag{2.2}$$

where $\beta$ is the lattice expansion factor ($cm^{-3}$) proportional to the covalent radius mismatch.

Dislocations have been known to affect dopant redistribution during thermal cycling by capturing and emitting point defects, leading to variations in junction depths. Calculations of the stress around a dislocation loop with observed density and size shows that the pressure around the loop layer can locally be on the order of $10^9$ dyne/$cm^2$. This value is comparable to stresses induced from patterned nitride films and isolation structures. Thus, the control and understanding of dislocation loops formation is critical.

The strain from standard CMOS unit processes, such as the shallow trench formation [23, 24], oxidation [25], silicide formation [26, 27], and contact etch stop layer (CESL) [28], has been observed to impact the transistor performance. Various modeling studies [29] and direct strain measurements from transmission electron microscopy (TEM) techniques have identified a change in channel strain from these processes. However, the stresses generated from each of these processes are misleading because the stress magnitudes are strong

**TABLE 2.2**
Material properties and calculated shear modulus (data from literature).

| | Young's modulus (GPa) | Shear modulus (GPa) | Poisson's ratio (-) | Coefficient of thermal expansion ($10^{-6}/^{o}$C) |
|---|---|---|---|---|
| Si | 190 | 77 | 0.23 | 2.6 |
| $CoSi_2$ | 160 | 61 | 0.3 | 10.4 |
| $Si_3N_4$ | 290 | 113 | 0.28 | 3 |
| $SiO_2$ | 60 | 24 | 0.25 | 3.3 |

functions of the device geometries and the mechanical constraints of the structure and thus quantitative strain estimation is difficult.

Silicidation (formation of silicides) is a thermal process used to form an alloy of silicon and refractory metal (e.g., Ti, Co, W, and Pt). The formation of silicides is typically accompanied by a net volume shrinkage and has been identified as a critical process step that may induce dislocations in the substrate (see Table 2.2). Etching steps also induce large topographical modifications in the microstructures. Continuous films generally produce only a very low level of stress. Problems occur when the surfaces of the films are not planar or contain discontinuities, such as window edges.

Silicides act as low-resistance contacts between the source, drain, and gate regions of a transistor. Ti, Co, and Ni are some of the common metals used to form silicides; based on the stoichiometry of the silicide compound that is formed and the thermal relaxation behavior, the magnitude of the strain introduced by these varies. In general, titanium silicides induce higher compressive strain than cobalt silicides [26]. Apart from the normal strains, silicides can also cause shear strain. Uneven growth interface results in shear components and stress concentrations, and this can be a source for dislocation nucleation in silicon if the silicidations are carried out at high temperatures [27].

In view of the aggressively scaled contact length, reduction of contact resistivity remains a challenge. In MOSFETs, source/drain resistance is very strongly influenced by contact resistance of the silicide layers, because it increases with the square of the linear scaling factor. Beyond the 90-nm node, the use of the currently used cobalt silicide ($CoSi_2$) contacts becomes unsustainable because of rapidly increasing resistance on narrow lines.

For nickel silicide formation, reduced thermal budget is needed partly because of the various phases of nickel silicides (NiSi, $Ni_2Si$, and $NiSi_2$) are forming at various temperature ranges. The most advantageous temperature range for the low resistance phase of NiSi is an initial anneal between 300 and 350$^{o}$C. This initial anneal is followed by etching of unreacted metal and a second activation anneal at a temperature of about 450$^{o}$C. The initial low temperature is needed for the NiSi phase formation to suppress Ni fast diffu-

sion which, if unchecked, leads to spike and encroachment effects.

NiSi or NiPtSi adoption is expected to enable the simultaneous achievement of targets for the contact junction depth, maximum silicon consumption, and contact silicide sheet resistance. NiSi silicide is also of interest for gate electrode application. It has been observed that dopant implantation into the poly prior to fully siliciding the polysilicon can allow achieving dual work functions needed for nMOS and pMOS transistors. In the longer term, a manufacturable and CMOS-integrable solution consisting of selective deposition of silicon or SiGe epitaxial layers is being sought. The adoption of epitaxially deposited SiGe, with its lower barrier height, could address both the contact resistivity and maximum silicide consumption issues. However, NiSiGe thermal stability is even lower than that of NiSi, thus further limiting the temperature of subsequent anneals to about 500$^o$C. Other materials such as platinum or platinum-nickel alloys are explored to improve thermal stability of the silicides.

Continuous films deposited over the entire silicon substrate produce low levels of stress in the silicon because the substrate is generally a few orders of magnitude thicker than the film. Large stresses occur when films are not planar or have discontinuities, and silicon substrate yields by forming dislocations. A good understanding of these effects and their interaction with the intentionally strained-Si layer is needed to ensure an optimum and stable process performance. In particular, a cobalt silicide ($CoSi_2$) film can be subjected to high residual stresses because of the vast difference between the CTE of this silicide and that of silicon [30]. Due to the existence of edge forces, discontinuities in the film introduced through the patterning give rise to localized stress fields in the substrate as well [31]. As VLSI technology is characterized by periodic material line patterns, it is important to understand how the different materials used and their patterning affects the stress always present in modern-day nanometer scale transistors.

Silicides are essential for good contacts with the device terminals. Since silicides are of different material than silicon, it is worth exploring the effect of silicide on the channel stress. Modern process technologies have nickel silicide (NiSi) for the contacts whereas formerly, cobalt silicide (CoSi), was used. Both the silicides have the same material properties (Young's modulus = 161 GPa, Poisson ratio = 0.33) [32]. Silicide thickness can range from 50-200 Å. The effect of thickness on strain of silicides is shown in Table 2.3. Since the silicide has material properties almost the same as that of silicon, its presence does not make any significant effect on the stress introduced into the channel.

During silicidation, the stress can arise from lattice mismatch between the substrate and film, the intrinsic stress related to the mechanical structure and properties of the film, and the difference between the thermal expansion coefficients of the film and the substrate. Table 2.4 lists the thermal expansion coefficients of several important silicides, together with those of the constituent metal and silicon. The thermal expansion coefficients of the

**TABLE 2.3**
Effect of Silicide on nMOS Channel Stress.

| Silicide thickness(Å) | $\sigma_{xx}$ (MPa) | $\sigma_{yy}$ (MPa) |
|---|---|---|
| 0 | 349 | -322 |
| 50 | 348 | -332 |
| 100 | 349 | -333 |
| 150 | 351 | -330 |
| 200 | 352 | -330 |

**TABLE 2.4**
Thermal properties of important silicides.

| Silicide | Melting point ($^oC$) | $\alpha$ (ppm/$^oC$) thermal expansion | Element | Melting point ($^oC$) | $\alpha$ (ppm/$^oC$) thermal expansion |
|---|---|---|---|---|---|
| $TiSi_2$ | 1500 | 12.5(200-1200) | Si | – | 3.0 |
| $CoSi_2$ | 1326 | 10.14(20-800) | Ti | 1661 | 8.5 |
| NiSi | 992 | | Co | 1495 | 12.0 |

silicides are considerably larger than those of the metals and silicon, and the difference can be responsible for the observed stress. The melting point of the transition metal silicides has also been listed in the table.

Deposited films such as $Si_3N_4$, $SiO_2$, TiN, TaN, Cu etc., are being used in back-end of the line. Nitride films are deposited on silicon after the silicide formation to act as the contact etch stop layer. Recent reports show the desired channel strain may be obtained for both the nMOS [33, 34] and pMOS [35]. The as-deposited CVD nitride films can introduce stresses greater than 1 GPa upon thermal treatment on planar silicon wafers, which arises from two sources: CTE mismatch between the silicon and nitride film and intrinsic film stress caused by film shrinkage. Based on the additives in the film, this stress can be engineered to be tensile or compressive [35]. When these films are conformally deposited on the nonplanar CMOS structure at the post-silicidation location, the actual strains that develop in silicon are more complicated.

The origin of intrinsic stresses comes from the energy configuration of the deposited atoms or ions that is often, just after the deposition, not in its lowest energy state. Compressive (or tensile) stresses develop in the layer when the atoms or molecules are closer (or further apart) in comparison to equilibrium. These intrinsic stresses may change during the post-deposition annealing because of atom/molecule reordering. These intrinsic stresses are generally thickness dependent. Processing conditions (i.e., temperature, pressure, deposition power, reactant and impurity concentrations, and type of material) are also important factors. The intrinsic stress is obtained by measuring the bending of the silicon wafer just after the deposition of the thin film.

**TABLE 2.5**
Mechanical properties and rheological behavior of several important materials used in microelectronic fabrication.

| Material | Young's modulus | Poisson's ratio | Thermal coefficient ($10^{-6o}C^{-1}$) | Intrinsic stress (GPa) |
|---|---|---|---|---|
| Oxide, Thermal | 66 | 0.17 | 0.5 | -400 |
| Oxide, HDP | 80 | 0.20 | 0.5 | - |
| Nitride, PECVD | 143 | 0.28 | 2.8 | - 50 |
| Nitride, LPCVD | 290 | 0.28 | 3.0 | 1300 |
| TEOS, LPCVD | 60 | 0.25 | 0.6 | - |
| Poly-Si, LPCVD | 180 | 0.27 | (2.8-4.5) | -350 |
| Ti, Sputtered | 115 | 0.34 | 8.6 | - |
| TiN, Sputtered | 410 | 0.30 | 9.6 | - |
| TaN, PVD | 457 | 0.25 | 6.5 | -1500 |
| W, CVD | 410 | 0.29 | 4.6 | 1000 |
| Cu, CVD | 87 | 0.34 | 16.6 | - |
| $CoSi_2$, Thermal | 160 | 0.25 | 10.8 | - |
| Co, Sputtered | 130 | 0.30 | 12.0 | - |
| Silicon | - | - | (2.8-4.5) | - |

Extrinsic stresses are caused by structural modifications in the film appearing after the deposition, and mostly due to further temperature changes or to interactions with the ambient. For example, silicon films undergo a density increase crystallization of the amorphous phase. Silicon dioxide exhibits densification phenomena corresponding to the removal of -OH bonds, formation of additional Si-O-Si bonds, and the rearrangement of these bonds leading to film shrinkage and tensile stresses. Mechanical properties and rheological behavior of several important materials used in microelectronic fabrication are shown in Table 2.5.

## 2.2 Global vs. Local Strain

At present, in CMOS transistor engineering, mainly two approaches are being used in obtaining the desired strain. One is based on developing the strain at the substrate level before the transistor is built. This is known as the "global" approach, for example, strained-Si on relaxed-SiGe virtual substrates. The approach depends largely on materials engineering, rather than device design. This approach stretches the silicon lattice by about 1%. For strained-Si, a graded layer of silicon germanium is grown on top of a bulk silicon wafer. A typically 2 $\mu$m thick SiGe layer has a 20% germanium concentration, with a higher concentration of germanium atoms at the top. Then a relatively thin

layer of silicon, about 20-nm thick, is placed on top of the SiGe layer. Over the past few years the technology to grow graded SiGe layers has matured. Strained-Si in which the SiGe layer is bonded to an insulating substrate, generally known as SiGe-on-insulator (SGOI) technology is now available. Strained silicon, while promising, faces several key challenges. Minimizing the number of dislocations within the silicon will be important to keeping yield rates high. Maintaining the level of strain during the manufacturing process is another challenge.

An alternate approach to utilize the epitaxial strain is to develop a compressive strain in the channel by SiGe deposition on the active regions beside the channel. In order to differentiate, this approach has been called the local-strain technology or, sometimes, the uniaxially stressed films. In the local-strain approach, the SiGe film is grown only in specific active regions after the transistor is formed by etching a recess in the silicon. In order to elucidate the application of the fundamentals discussed in the initial sections, we discuss this approach in detail.

"Local" strained-Si technology is incorporated during the transistor fabrication process. The strains induced during transistor processing are typically uniaxial (i.e., in one direction) and are incorporated via tensile/compressive capping layers or recessed epitaxial film deposition in the source-drain regions. These processes are generally not universal in their implementation and need to be tailored to a particular transistor integration scheme. The straining technique based on process is known as process-induced strain where stress is induced in a specified zone or "local" in the transistor. However, the scalability of local strain is one of the most important topics for future CMOS performance. A combination of global and local approaches is also feasible and presently under active investigation [36]. In principle, three kinds of stresses can be applied to the transistor:

- a uniaxial stress along one crystallographic direction,
- a biaxial stress along two perpendicular crystallographic directions, and
- a hydrostatic stress applied uniformly in all three directions.

Ge et al. [22] have demonstrated a process-strained Si (PSS) CMOS technology using the concept of 3D strain engineering. Methods of producing PSS include stress engineering of trench isolation, silicide, and cap layer, to improve nMOS and pMOS performance simultaneously. Each of these approaches results in a 5-10% enhancement in the ring oscillator speed. It has been shown that by increasing tensile strain (or decreasing compressive strain) in y direction benefits nMOS and pMOS simultaneously while strain changes in the other two directions result in nMOS and pMOS performance trade-off. It is therefore desirable to have a predominant tensile strain in the y direction for overall CMOS performance improvement. Fig. 2.3 illustrates three of these techniques, With optimized stress engineering in the cap-layer, STI and silicidation processes, both nMOS and pMOS improvements up to 15% are obtained in a typical 0.13 $\mu$m technology (see Fig. 2.4).

Since all these process-induced strains in the channel are a function of the

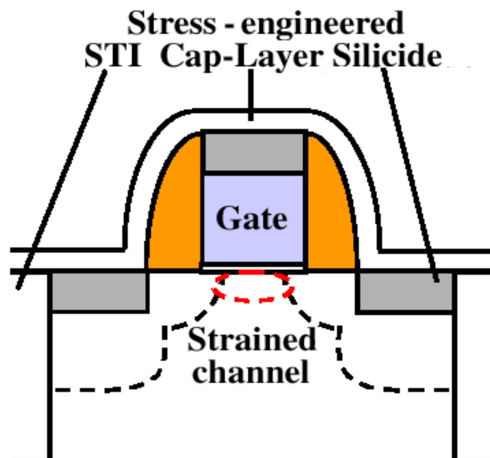



**FIGURE 2.3**
Schematic of device structure featuring process strains induced by stress-engineered trench isolation, cap-layer and silicide. After C.-H. Ge et al., *IEEE IEDM Tech. Dig.*, 2003(73-76). With permission.

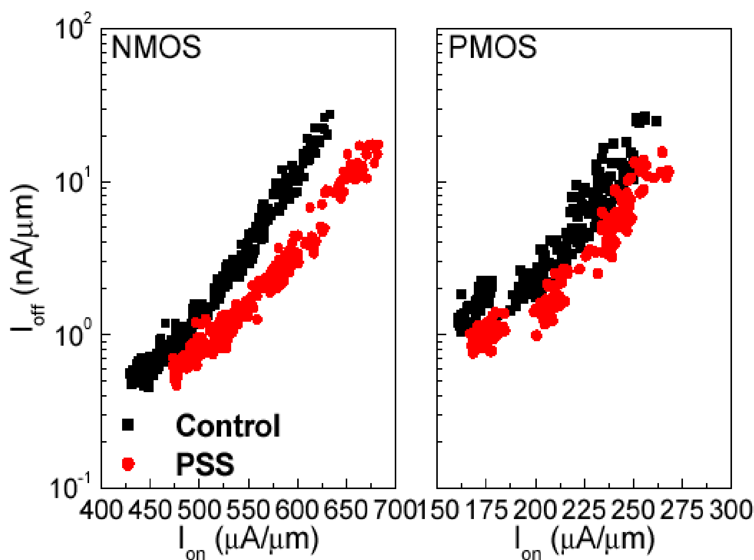



**FIGURE 2.4**
$I_{on}$ vs. $I_{off}$ characteristics at 1 V operating voltage. Up to 15% improvement is achieved in nMOS and pMOS with cap-layer, STI- and silicide-strained Si. After C.-H. Ge et al., *IEEE IEDM Tech. Dig.*, 2003(73-76). With permission.

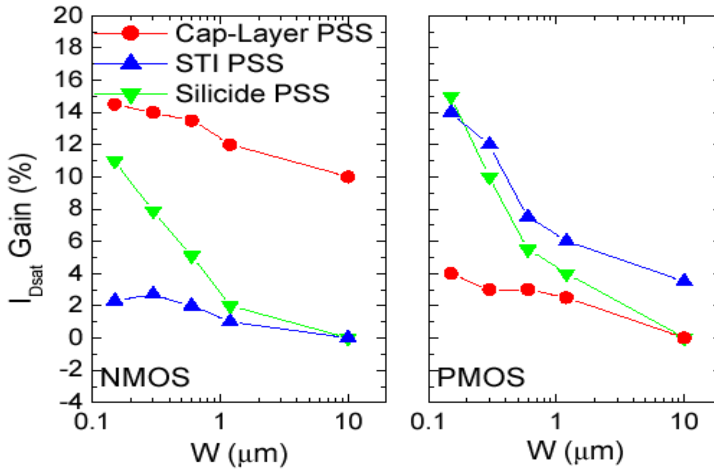



**FIGURE 2.5**
Both improved nMOS and pMOS performance at reduced channel width confirms increased tensile strain (or reduced compressive strain) in y direction with the three PSS techniques. L = 90 nm. After C.-H. Ge et al., *IEEE IEDM Tech. Dig.*, 2003(73-76). With permission.

contact area between the stressor and source/drain as well as the distance between the stressor and the channel center, they have various degrees of dependence on channel length as well as channel width. As evidenced in Fig. 2.4, device performance increases with higher current drive and therefore with shorter channel length due to increased channel strain. Width dependence of CMOS performance enhancement using each of these process-strain techniques is illustrated in Fig. 2.5 at a physical gate length of 90 nm. It is evident that both nMOS and pMOS performance improves with reduced channel width for all the three PSS techniques due to increased tensile strain in the y direction. Note that STI- and silicide-induced strains in the channel seem to show more dependence on device width than cap-layer induced strain. This may imply that the stress engineering of STI and silicidation processes are more effective in increasing tensile strain (or decreasing compressive strain) in y direction.

## 2.3 Substrate-Induced Strain

A widely adopted method to introduce wafer-based biaxial stress to enhance CMOS performance is practiced by growing a silicon film atop relaxed-SiGe

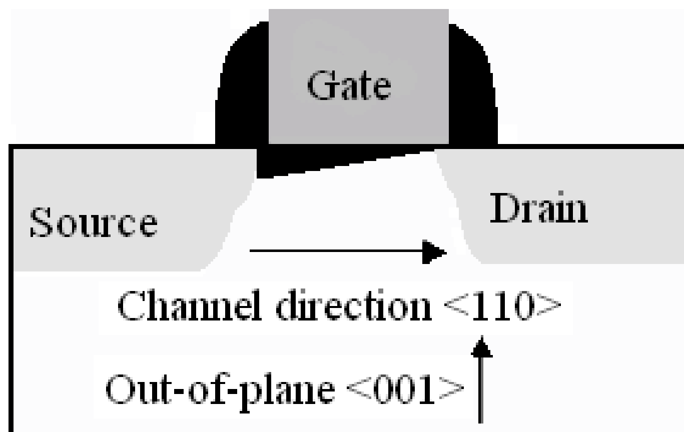



**FIGURE 2.6**
MOSFET schematic device cross-section.

virtual substrate. Due to the lattice mismatch between Si and Ge atoms, tensile biaxial stress is generated in Si film, which enhances the performance of nMOS and pMOS. It can be seen from the piezoresistance coefficients of Si for standard layout and wafer orientation, that nMOS performance (electron mobility) is enhanced by uniaxial longitudinal tensile and out-of-plane compressive stress (Fig. 2.6), while pMOS performance (hole mobility) is enhanced by uniaxial longitudinal compressive stress. Interestingly, both nMOS and pMOS performance is enhanced by biaxial tensile stress.

Global strained-Si technology is based upon the deposition of $Si_{1-x}Ge_x$ alloys, which form a template for subsequent channel layer deposition (and layer transfer process in the case of strained-Si on insulator). Typically the alloys contain a Ge content in the range of x = 0.15 to 0.30. The deposition of pure Si on such a template results in biaxial (i.e., in the two in-plane directions) strains on the order of 1%. Due to strain introduction at the starting materials stage, global or wafer-scale strained-Si is more universal in nature and some of its attributes or specifications can be addressed more generally.

The benefits offered by wafer-level strain are highly desired by the IC industry. However, a major concern about the manufacturability of this technology has been its high level of crystal defects in the wafer-level strained silicon layer - typically resulting from the epitaxial technique used to grow the silicon germanium templates that produce the strain in the top silicon layer. Generally, three areas of development that need to be considered when considering the commercial use of global strained-Si technology are:

- Concurrent optimization of nMOS and pMOS transistors needs to be

explored more thoroughly. A systematic development effort is required to elucidate the complete benefits of global strained-Si technology,

• Defects (misfit dislocations) at gross levels at the interface between strained-Si and the $Si_{1-x}Ge_x$ alloy have been shown to increase transistor leakage levels. However, no significant results have been reported on the mechanisms for the leakage or methods to mitigate such leakage. Further research and development is required to improve the understanding on this topic, and

• Leakage and reliability mechanisms associated with threading dislocations and other defects are not yet clearly understood.

## 2.4 Process-Induced Stress

Process-induced strain refers to the strain in silicon lattice generated by many of the standard CMOS processing steps, some of which can be engineered to enhance/reduce strain. Many recent publications on production-worthy CMOS technologies reported about engineering the process-induced strain to derive transistor performance improvements [37]. An alternate approach to utilize the epitaxial strain is to develop a compressive strain in the channel by SiGe deposition on the active regions beside the channel [18, 19]. A strategy to increase compressive stress generated by embedded SiGe source/drain regions in a pMOS device have been studied. Uniaxial front-end process-induced stress is being widely adopted in almost all logic technologies because of its edge over the biaxial counterpart [1, 21]. The topology of the source/drain regions can assist to increase the stress and this property is well exploited during the front-end processing. Raising the SiGe source/drain regions up to a certain extent transfers higher stress to the channel, thereby further improving the mobility of holes. Increasing the raised step height an increase in stress is obtained. However, beyond approximately 300 Å of step height, the raised SiGe source/drain regions do not increase the stress further. Higher stress may be induced by increasing the Ge concentration. Higher Ge concentration in SiGe compresses the silicon lattice even greater, which is simulated by higher intrinsic stress inside SiGe regions. In state-of-the-art nanodevice research, Ge concentration is increased up to 30%.

The concept of improving electron mobility in n-channel devices by depositing a tensile $Si_3N_4$ capping layer over the entire structure has been demonstrated [35]. It has been shown that by increasing the $Si_3N_4$ film thickness the stress in the channel may be increased. This holds true up to a certain thickness in a legged gate structure, which are widely used in logic technologies. The effect of increasing the nitride thickness, keeping a constant space between two nMOS devices in stacked gate structure, has been studied. Fig. 2.7 shows horizontal |xx| and vertical |yy| strain at center of the channel re-

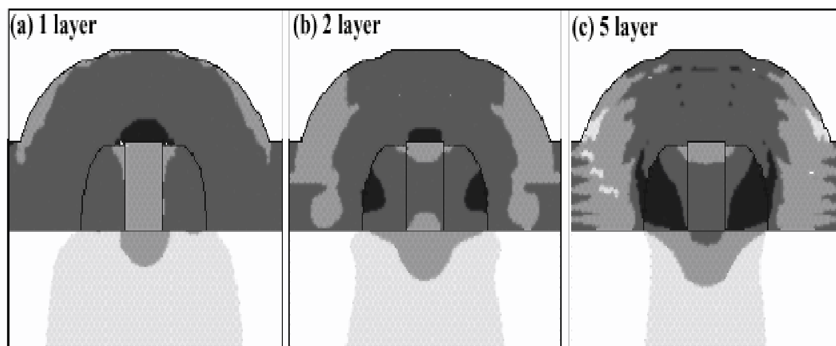



**FIGURE 2.7**
2D Stress simulation results (vertical strain yy) of n-MOSFETs with (a) 1 layer (b) 2 layer and (c) 5 layer SiN tensile films. Total SiN thickness is 100 nm in all devices. Strong strain is observed at sidewall of (c) 5 layer. After K. Goto et al., *IEEE IEDM Tech. Dig.*, 2004(209-212). With permission.

gion by changing the total SiN thickness and layer number. Although strain of single layer becomes saturate when SiN thickness is more than 100 nm, double and five SiN layers shift the saturation point more than 140-nm. Stress gain by changing the layer number is calculated and is shown in Fig. 2.8. One observes that multilayer structure is effective for thicker SiN. For 140-nm thick SiN case, strain is increased up to 28% and 55% by using double and five SiN layers, respectively.

Due to the inherent mechanical properties of silicon, the stress field generated by a stress source (such as the STI sidewall) penetrates through approximately 200 nm of adjacent silicon. Stress decays inversely proportional to distance from the stress source and drops by a factor of three within $\sim$ 200 nm. When several stress sources happen to be within that range, their stress fields overlap. This is why devices at the 180-nm technology node and below have significant stress levels due to the overlapping stresses from STI, gate stack, and silicide.

Table 2.6 summarizes the impact of 1 GPa stress applied along x, y, and z in the channel on the transistor performance. It is based on a simple classic piezoresistance effect [7], yet provides reasonably good description of the stressed devices [21]. The performance modification percentages listed in the table are only approximate, as they reflect the changes in the low field mobility and neglect a variety of other factors. Compressive and tensile stresses cause opposite effects on transistor performance.

The process-induced straining techniques discussed above are standard steps in the CMOS flows. There are also reports on deliberate disposable film deposition to generate strained channels. Ota et al. [33] reported significant

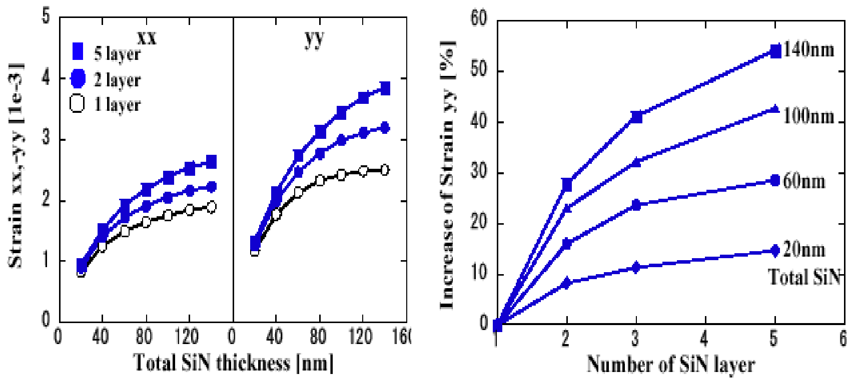



**FIGURE 2.8**
Horizontal xx, and vertical yy strain under the gate varying total SiN thickness. Though stress of single layer saturates over 100 nm, multi (2 and 5) layer keeps increasing. Increasing ratio of channel strain |yy| with varying SiN layer. Multilayer SiN capping is very effective for thicker SiN. After K. Goto et al., *IEEE IEDM Tech. Dig.*, 2004(209-212). With permission.

**TABLE 2.6**
Impact of 1 GPa stress in Si on electron and hole mobility enhancement. Computation is based on piezoelectric effect. After V. Moroz et al., *Solid State Technol.*, July 2004.

| Stress component | Tensile | Tensile | Compressive | Compressive |
|---|---|---|---|---|
| | nMOS | pMOS | nMOS | pMOS |
| 1 GPa along channel (x) | 30% | -70% | -30% | 70% |
| 1 GPa across channel (z) | 20% | 70% | -20% | -70% |
| 1 GPa vertical (y) | -50% | 1% | 50% | -1% |

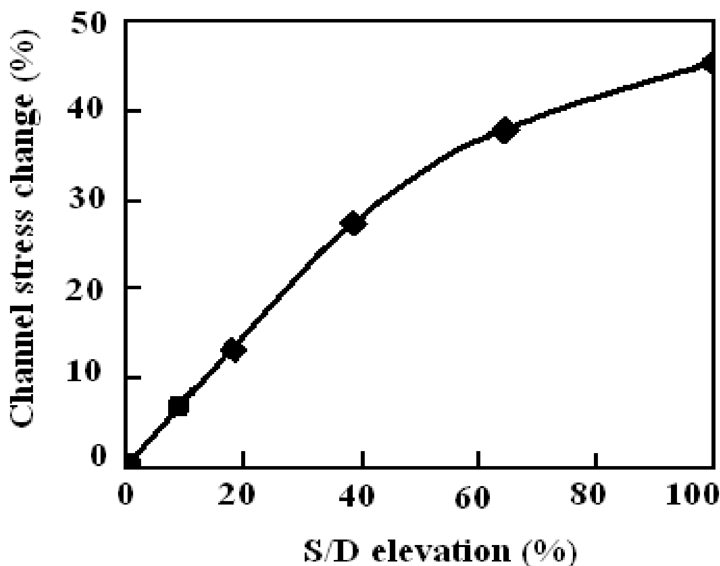



**FIGURE 2.9**
Sensitivity of channel stress to S/D elevation. After F. Nouri et al., *IEEE IEDM Tech. Dig.*, 2004(1055-1058). With permission.

strain-related nMOS performance improvement from stress-memorization effect. They deposited a tensile stressed oxide film on amorphized polysilicon and annealed the device at high temperature and stripped the stressed film. The observed transistor improvement was attributed to the channel strain that was locked in place by the polysilicon regrowth process. Although not very well understood, the strain induced by this disposable film continues to play a significant role in nMOS-performance engineering [38].

When a stress is introduced, the lattice structure is deformed in response to the superposition of a biaxial tensile strain and a smaller uniaxial tensile or compressive stress. On atomistic level, stress causes strain, which is the displacement of atoms from equilibrium lattice positions. The change in lattice spacing changes the effective mass of the carriers and affects the scattering rates they undergo in the crystal lattice. Uniaxial strain has shown superior pMOS mobility enhancement but has been only available as local strain using techniques such as stressed capping layers, or source/drain selective SiGe epi deposition such as Intel's 90-nm process currently in production [1].

Sensitivities of the channel stress to the S/D elevation and to gate-STI distance have been investigated by Nouri et al. [39]. Fig. 2.9 shows that with the increase of S/D elevation, the stress increases linearly up to 40% and then saturates. Transistors with shallow SiGe S/D and short gate-to-STI distance exhibit high stress sensitivity to the layout variations.

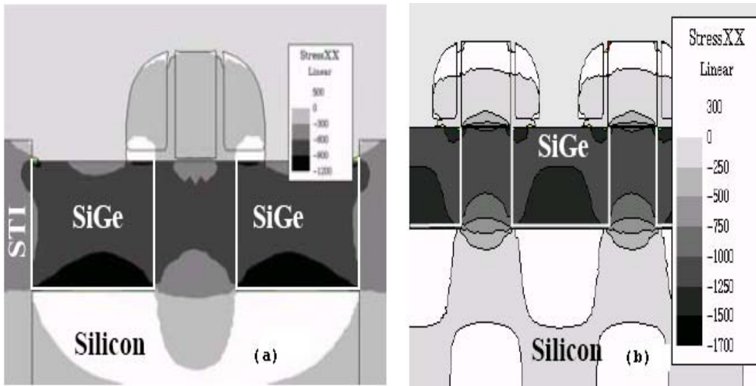



**FIGURE 2.10**
(a) Simulated structure for the isolated transistor, showing stress in the channel direction (MPa) and (b) Simulated structure for the dense transistors, showing stress in the channel direction (MPa). After G. Eneman et al., *IEEE VLSI Tech. Symp. Dig.*, 2005(22-23). With permission.

The results of a study on the layout dependence of a SiGe S/D p-MOSFET technology have been reported by G. Eneman et al. [40]. Simulated structures with stress contour for the isolated and dense transistors are shown in Figs. 2.10(a) and (b). The stress in the channel as a function of the source-drain length is shown in Fig. 2.11 for a recess depth of 30, 70, and 120 nm. While for long $L_{s/d}$s, stress does not depend on the source-drain length, for $L_{s/d}$ below 500 nm, the dependence becomes significant, especially for larger S/D recess. The stress reduction, when $L_{s/d}$ decreases from 1.5 $\mu$m to 500nm is 17%, 11%, and 4% for recess depths of 120, 70, and 30 nm, respectively.

For the simulated dense transistor structure (see Fig. 2.10(b)), the volume of SiGe present is larger than for isolated transistor structures, leading to higher stress values, and faster saturation of the stress levels for large $L_{p/p}$s (see Fig. 2.12). The stress reduction when $L_{p/p}$ is found to decrease from 1.5 $\mu$m to 500 nm is 9%, 6%, and 4% for etch depths of 120, 70, and 30 nm, respectively.

Biaxial tensile strained-Si is considered as a technique in the last decade. However, it has been difficult to implement because of misfit and threading dislocations, Ge outdiffusion, silicide formation difficulty, self-heating effect, higher As diffusion in the S/D extension region, and cost. DSL (contact etch stop layer) provides enhancement in both nMOS and pMOS. It needs extra steps and contact opening may be a concern.

Local strain can be applied to the channel through a Stress Memorization Technique (SMT). In the conventional fabrication process, Source/Drain Si area and poly-gate are amorphized by S/D and extension implantation. In the stress memorization process, conventional dopant activation spike anneal

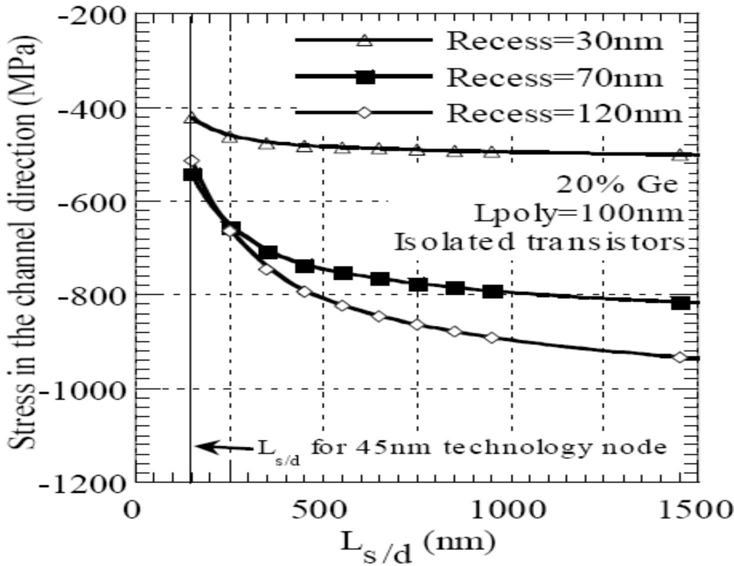



**FIGURE 2.11**
Simulation of stress in the center of the channel as a function of gate to active edge distance for isolated transistors. After G. Eneman et al., *IEEE VLSI Tech. Symp. Dig.*, 2005(22-23). With permission.

is performed after the deposition of a tensile stressor capping layer, such as nitride [33, 37]. The nitride layer is subsequently removed, and the next step is salicidation. The stress effect is transferred from the nitride film to the channel during annealing and memorized by the recrystallization of the S/D and poly-gate amorphized layers. After active area and gate recrystallization, the stress inside the channel is memorized. Since the nitride film is disposable, a very thick capping layer can be used to increase the stress level without any process limitation. However, SMT needs extra steps although it is easy to implement. It does not have any major process or ground rule limitation. However, it can only improve nMOS performance. Strained-SiGe in S/D structure provides highly compressive strain to the channel, and results in large hole mobility improvements. Extra complicated integration steps are necessary; nevertheless, it can provide very good pMOS performance. Hybrid orientation technology (<110> substrate) clearly shows an attractive option to improve device performance. However, the impact on circuit performance of mixing SOI and bulk devices on the same chip and circuit design will require more detailed analysis. Channel orientation (<100> channel, 45 deg notch) is a highly manufacturable technique to reduce integration complexity. It allows nMOS improvement by another stress transfer process, without any pMOS degradation. However, pMOS performance cannot be further improved by other stress transfer method. Some key differences between wafer-level

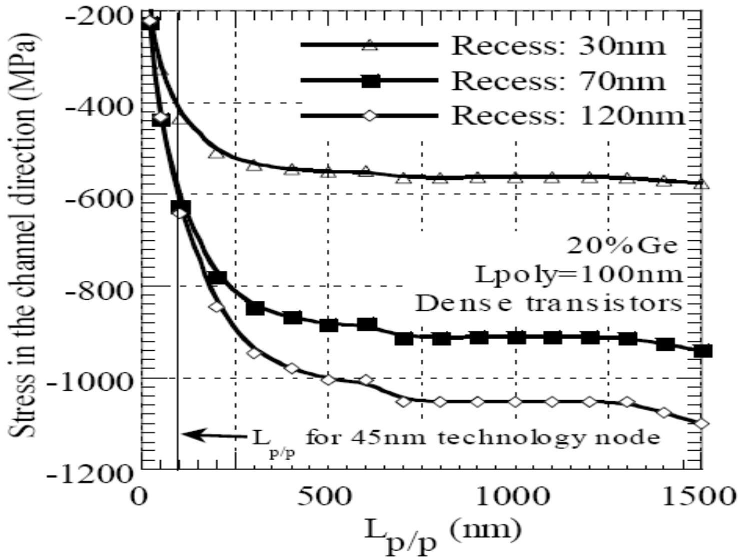



**FIGURE 2.12**
Simulation of stress in the center of the channel as a function of poly-to-poly distance for dense transistor structures. After G. Eneman et al., *IEEE VLSI Tech. Symp. Dig.*, 2005(22-23). With permission.

process-induced uniaxial strain and biaxial SiGe-based strain are as follows:

- Uniaxial strain can be applied along a single direction on a chosen surface orientation: e.g., along <110> direction on a (100) surface.
- Tensile or compressive strain can be achieved with direct mechanical stress. Only tensile biaxial strain has been reported with SiGe-based strain.
- Uniaxial strain can utilize hybrid orientation process technologies to selectively process nMOS and pMOS devices efficiently.
- Uniaxial strain offers significant pMOS device performance advantages over biaxial strain at realistic gate electric fields ($\sim$ 1 MV/cm).
- Direct mechanical stress technology can be used to generate biaxially strained layers with independent control of strain along x- and y-axes.
- Although both SiGe-based and uniaxial SSOI uses layer-transfer, the direct mechanical technique does not require a relatively expensive relaxed-SiGe donor.

### 2.4.1 Mechanical Stress Control

Mechanical stress control in the channel regions of MOSFETs enables one to overcome the limitations on the scaling down of devices. A recent study has shown that mechanical stress from a contact etch-stop silicon nitride (SiN) layer affects the drive current [41, 42]. Changing the Si lattice spacing to a

value other than the equilibrium value by using mechanical stress increases the mobility of electrons and holes. However, when the mechanical tensile stress of the SiN layer increases, n-channel drive current increases while the p-channel drive current decreases. However, when the stress becomes compressive, the opposite occurs. The n- and p-channel drive currents are, therefore, in a trade-off relationship in terms of mechanical stress. Local strained channel technique that uses a capping SiN layer and stacked a-Si/polySi gate structure or local straining technique using strained polySi gate electrode [33] also significantly improves the current drivability by controlling the thickness of stacked a-Si and capping SiN-layer, without any degradation on the interface at oxide/Si-substrate, compared to that of the conventional single polySi gate device.

Mechanical stress can be transferred to the channel through Si active area and poly-gate if a permanent stressed liner is deposited on a device. Tensile liner will enhance nMOS and compressive liner will enhance pMOS mobility, respectively. After silicide formation (either cobalt or nickel silicide), a highly stressed liner (either tensile or compressive), such as plasma-enhanced chemical vapor deposition (PECVD) or rapid thermal CVD nitride, is uniformly deposited over the wafer. The drive current dependence on stress is through the change in film thickness and material. In 2001, researchers from Hitachi developed a stressing technique, called "local mechanical-stress control" (LMC), to enhance the CMOS current drivability [4]. It utilizes heavy mechanical stress produced by a SiN layer and Ge-ion implantation to selectively relax the stress of the layer. The drive currents of both n- and p-channel MOSFETs could be improved by controlling the stress of the SiN layer selectively without degrading their reliability. The effects of LMC become more significant as devices become smaller, and the drive current is estimated to increase by more than 20% in the future 70-nm CMOS technology node.

The authors tested two LMC techniques (LMC (A) and (B)) on two sets of CMOS devices fabricated using conventional salicide processes (Fig. 2.13). The LMC technique includes two additional steps. The process sequence (Fig. 2.13(a)) for LMC (A) was as follows: First, a SiN layer with highly compressive stress was deposited over both an n- and a p-channel MOSFET, using plasma-enhanced CVD. Then, Ge ions were selectively implanted into this layer on the n-channel MOSFET (Fig. 2.13(b)). Fig. 2.14 shows the cross-sectional TEM photograph of typical n-channel MOSFET and strain distribution at Si surface near the Si-$SiO_2$ and Si-Silicide interface. The strain was estimated from the change of Si lattice spacing obtained by focused electron beam diffraction patterns. This technique has a resolution of $\sim$ 10 nm in diameter. The effects of LMC (A) and (B) on the $I_{on}$-$I_{off}$ characteristics and the dependence of $I_{on}$ at fixed $I_{off}$ on the LMC (A) and (B) conditions are shown in Fig. 2.15(a) and (b), respectively.

Bera et al. [43] have reported a novel post-backend strain technique and its impact on MOSFET device performance. By bonding the Si wafer after transistor fabrication onto a plastic substrate (a conventional packaging material FR-4), a biaxial-tensile strain ($\sim 0.026\%$) was achieved globally and uniformly

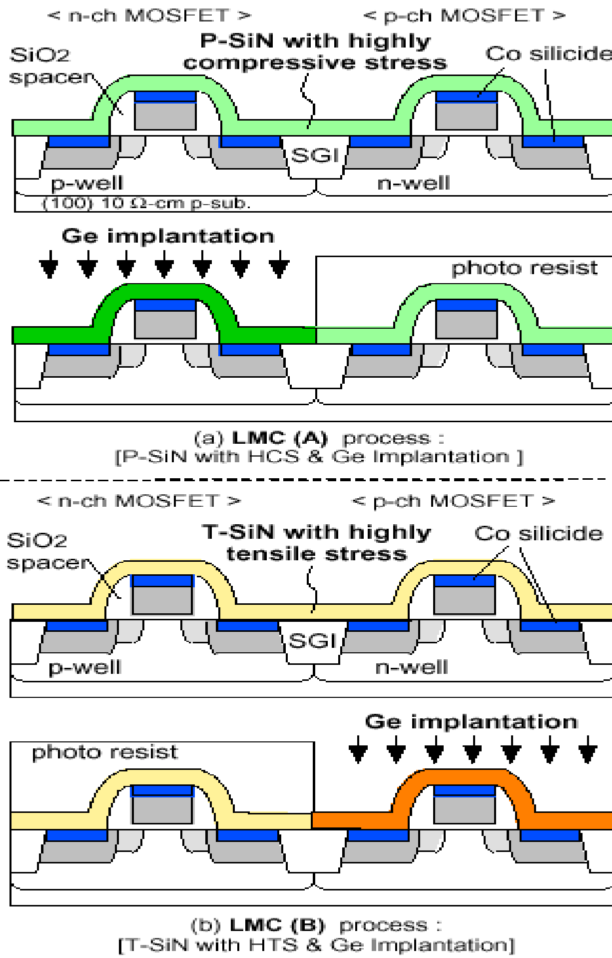



**FIGURE 2.13**
Fabrication process sequences for local mechanical stress control. After A. Shimizu et al., *IEEE IEDM Tech. Dig.*, 2001(433-436). With permission.

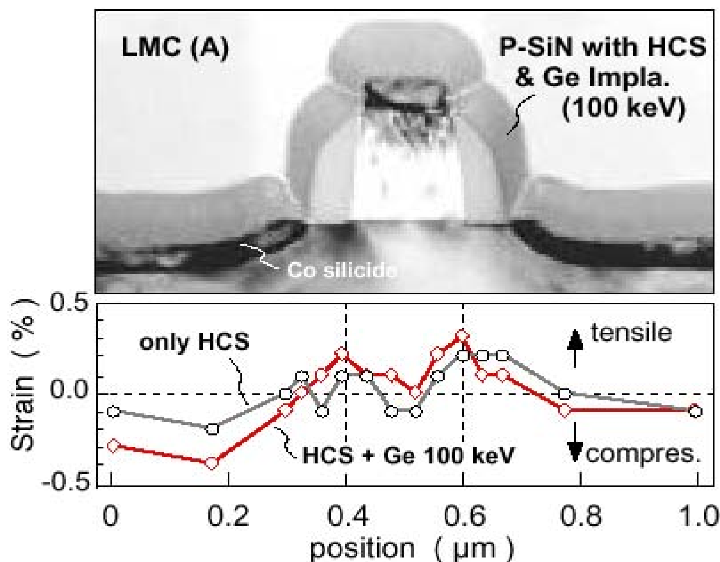



**FIGURE 2.14**
Cross-sectional TEM photograph of typical n-channel MOSFET and strain distribution at Si surface near the Si-$SiO_2$ and Si-silicide interface. After A. Shimizu et al., *IEEE IEDM Tech. Dig.*, 2001(433-436). With permission.

across the wafer due to the shrinkage of the bonded adhesive. A drain current improvement (average $\Delta I_d/I_d \sim 10\%$) for n-MOSFETs uniformly across the 8-in wafer is observed, independent of the gate dimensions ($L_g \sim 55$ nm to 0.530 $\mu$m/W $\sim$ 2 to 20 $\mu$m). The p-MOSFETs also exhibited $I_d$-improvement by $\sim$ 7% under the same biaxial tensile strain.

### 2.4.2 Substrate Orientation Dependence

Mobility enhancement is an attractive option because it can potentially improve device performance beyond any of the benefits resulting from device scaling. The two main approaches, as discussed above, being pursued are strain engineering (both process- and substrate-induced) and orientation effects. Strain effects induced during the fabrication process can increase the channel mobility. Both tensile and compressive stresses can be introduced in any one of three dimensions by process techniques. The electron and hole mobilities have different dependencies on the type of uniaxial stress.

Effects of biaxial and uniaxial strain on carrier mobility in bulk- and ultrathin-body MOSFETs have been systematically investigated using externally applied mechanical stress [44]. It has been demonstrated that

(a) in bulk n-MOSFETs, electron mobility enhancement is stronger in the following order of biaxial tensile, <100> uniaxial tensile, and <110> uniaxial

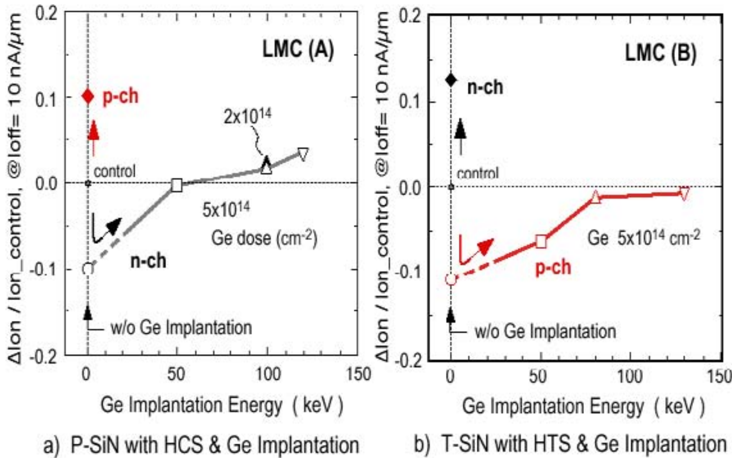



**FIGURE 2.15**
Change of drive current in LMC (A) and (B) technique. The $I_{on}$ that degraded due to highly compressive stress (or highly tensile stress) of SiN was recovered by the LMC to the control value. $I_{on}$ of n-channel in LMC (A) was even larger than in the control process, because the SiN in the control process had small compressive stress. After A. Shimizu et al., *IEEE IEDM Tech. Dig.*, 2001(433-436). With permission.

tensile strains, and

(b) in bulk p-MOSFETs, hole mobility enhancement is stronger in the order of <110> uniaxial compressive, <100> uniaxial compressive, and biaxial tensile strains.

It has also been demonstrated that uniaxial strain enhances both the electron and hole mobility in UTB MOSFETs with thin SOI of at least 2.5 nm. Biaxial tensile strain enhances only the electron mobility in UTB MOSFETs with thin SOI of less than 5 nm [45]. It is also predicted that the strain engineering along with subband structure engineering in UTB MOSFETs can further enhance the mobility.

Inversion layer mobility depends on surface orientations and current flow directions. For p-channel MOSFETs, hole mobility is 2.5 times higher on (110)-oriented surfaces compared with that on standard wafers with (100) surface orientation. However, electron mobility is the highest on (100) substrates. To realize the advantage of carrier mobility dependence on surface orientation fully, a new technology to fabricate CMOS on hybrid substrates with different crystal orientations has been developed, with nMOS on Si with (100) surface orientation and pMOS on (110) surface orientation. It has been shown that the strain and orientation effects on transistor performance are additive. Full CMOS integration with ring-oscillator delay reduction has been demonstrated. Strained-Si channels can also be produced by depositing Si

epitaxially on relaxed-SiGe crystals. The strain is created from the differences in lattice spacing between Si and SiGe. While transistor performance enhancement has been reported, its progress has been hindered by the lack of low-cost, high-quality strained-Si substrates, as well as the challenges in materials and process integration. Alternative high-mobility substrates, such as the use of Ge and III-V compound materials, are being considered for post-CMOS applications. Basic scaling elements such as gate dielectrics and junction technology in these substrates are still in their very early research stages.

### 2.4.3 Strained-Ge

Transistors with pure Ge channels are being investigated for implementation in next-generation microelectronics. The ultimate advantage of working with Ge is the carrier mobility enhancement, making it attractive for high-speed circuit applications. Low field electron mobility in Ge is more than double that of Si (3900 vs. 1500 $cm^2/V$-s) and the increase is four-fold for holes (1900 vs. 450 $cm^2/V$-s). Despite the intrinsic speed advantages of implementing Ge transistor technology, Ge has not established a strong presence as an electronic material for ubiquitous microelectronic application because it does not form a stable oxide, which is critical for gate electrode formation, limiting its utility in traditional MOS manufacturing methods. However, as the industry is considering a transition to high-k dielectric films, necessitated by electron tunneling through very thin $SiO_2$, opportunities are being explored to implement pure Ge transistor technology. There are two aspects of development that need to be considered in parallel for implementation of Ge transistor technology: (a) the embodiment of the Ge substrate technology and (b) the multiple aspects of Ge transistor integration and performance optimization.

## 2.5 Stress/Strain Analysis

As active area size shrinks, the stress of the gate, spacers, and silicides have a significant impact on the stress state in the transistor channel and, thereby, can have a strong influence on the device characteristics. Stress modeling of front-end processes has been used in the past to understand materials processing issues. One of the first applications was to simulate stress-dependent oxide growth in 2D and 3D using either viscous or viscoelastic material models [46, 47, 48, 49]. Simulation of the stress in silicon from a variety of strain sources including oxidation, thermal mismatch, intrinsic stress, and dopants have also been presented [50]. Stress simulations have been used to understand the stress effects on electrical characteristics of MOSFETs. Stress from

shallow trench isolation was used to explain reduced electron mobility due to layout [51]. Smeys et al. [11] used LOCOS stress simulations to explain increased junction leakage current.

Process stress simulations using FLOOPS [52] have been reported by Cea et al. [53]. The basic equation solved for stress simulations is the balance of forces given by:

$$\sum_{j=x,y,z} \frac{\partial \sigma_{ij}}{\partial j} + F_i = 0, (for\ i = x, y, z) \tag{2.3}$$

where $\sigma$ represents the stress and $F_i$ represents the total external forces. The types of external forces that are important for stress simulations include intrinsic film stress, thermal strain, misfit strain due to doping or epitaxy, and stress due to material growth. Eqn. 2.3 is solved using the finite element method using linear elements: triangles in two-dimension and tetrahedra in three-dimension. The plane strain approximation is used in 2D. FLOOPS uses a viscoelastic mechanical model given in one-dimension by

$$\dot{\epsilon} = \frac{\dot{\sigma}}{E^o} + \frac{\sigma}{\eta} \tag{2.4}$$

where $\sigma$ is the stress, $\dot{\epsilon}$ is the strain rate, $E^o$ is Young's modulus and $\eta$ is the viscosity, which can be stress dependent.

In a typical process flow without stress engineering, stress-field generation is primarily due to the different volume expansion/contraction rates of the adjacent materials during temperature ramps. Stress also is generated by growing oxides and silicides, etching and depositing layers, and introducing dopants into the silicon. Stress can be introduced intentionally by deposition of a pre-strained layer or it could be due to unintentional stress sources. Unintentional stress is generated as a side effect of oxidation, etch, deposition, silicidation, and thermal steps. The unintentional stress should be considered to determine the overall stress distribution when designing appropriate methods to introduce intentional stress into the device. While it is difficult to measure stress distribution in submicron devices, 2D and 3D stress distribution and its evolution during the process flow can be simulated.

Moroz et al. [54] have explored stress effects in the 45-nm MOSFET with the Taurus-Process simulator that models stresses for semiconductor applications. For a MOSFET with pre-strain deposited layers with 1 GPa intrinsic stress, the average stress in the channel is shown in Fig. 2.16(a). Fig. 2.16(b) shows the distribution of the x stress component in the 45-nm transistor with 15% Ge in SiGe source/drain (S/D) at the end of the process flow. A uniform compressive stress of $\sim$ 1 GPa is generated in the channel. Some of the methods possess relatively high stress-transfer efficiency. Notably, SiGe S/D and strained STI transfer slightly more than 50% of their stress into the channel. This helps to introduce significant stress into the channel, while keeping the peak stress low to prevent defect/crack formation. It is clear that all methods

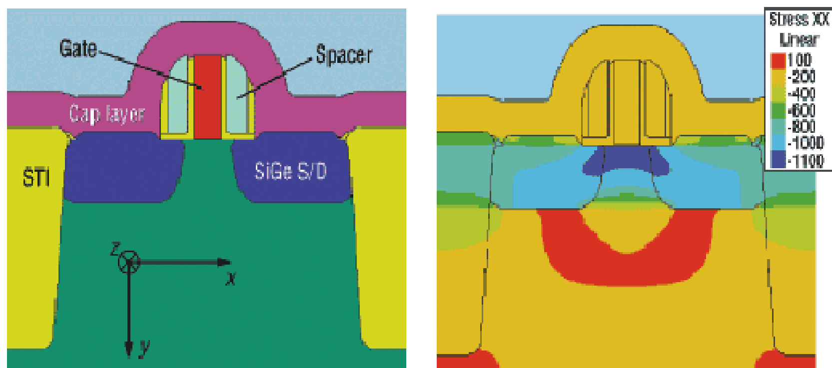



**FIGURE 2.16**
(a) Cross-section of a 45-nm CMOS transistor with SiGe S/D. (b) Distribution of the x stress component (i.e., in the direction along the channel) in a 45-nm MOSFET at the end of the process flow. After V. Moroz et al., *Solid State Technol.*, July 2004.

of introducing stress have an efficiency (ratio of stress in channel to intrinsic stress) $<$ 100% (see Fig. 2.17). Positive numbers mean that the stress in the channel has the same sign as the stress in the deposited layer. Negative percentage means that the stress in the channel is opposite to the original stress in the deposited layer.

It is clear that none of these methods creates a purely uniaxial stress, but some of them create stress patterns with one dominant component and therefore come close to uniaxial. It is important to make sure that the combined effect of the three stress components (along x, y, and z axes) is still positive and different contributions do not cancel each other out. If several methods are used simultaneously, then certain stress components add up, but some others do not, exhibiting very nonlinear interaction. For a different device geometry, the stress transfer pattern can be completely different. For example, Fig. 2.18 shows the 3D stress pattern for a 25-nm fully depleted SOI MOSFET with metal gate and mesa isolation. In this case, a complicated 3D stress pattern is formed with significant x, y, and z components. The reported electrical measurements of such nMOS and pMOS devices [55] show that stress enhanced performance of both nMOS and pMOS by $\sim$ 30%. Fig. 2.18 shows that the distribution of the z stress component is very nonuniform in the channel, varying from 100 MPa compressive to 800 MPa tensile stress. Stress transfer efficiency from metal gate to the channel is close to 100%. This must be attributed to the specific geometry of this device.

The summary of stress-engineering effects is presented in Table 2.7. It is based on piezoresistance model and applies to the particular case of geometry and process flow of 45-nm MOSFET transistors explored in this work. Different geometries, process flows, and technology nodes will yield different

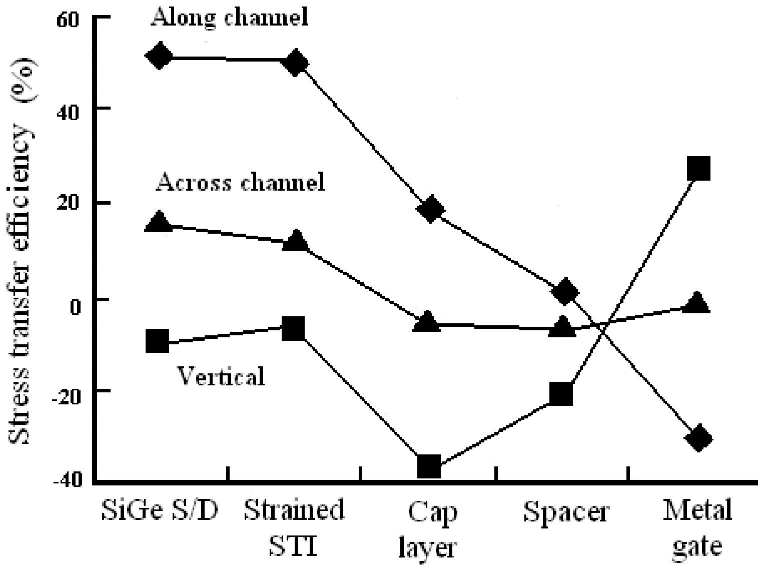



**FIGURE 2.17**
Efficiency of transferring stress from different sources into the channel. After V. Moroz et al., *Solid State Technol.*, July 2004.

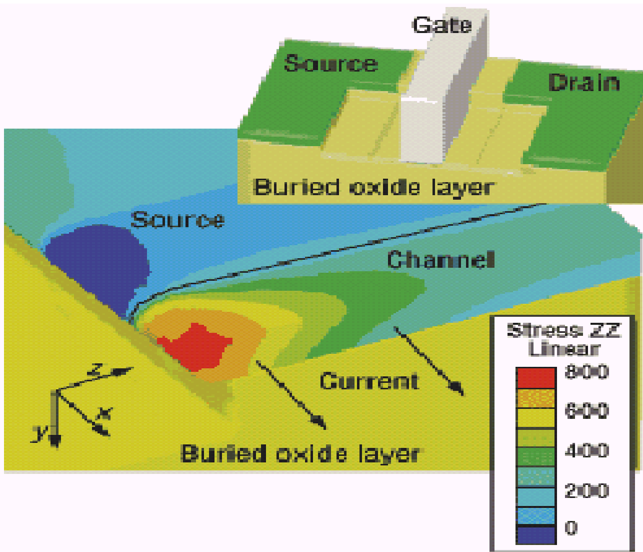



**FIGURE 2.18**
Distribution of z stress component in 25-nm FDSOI MOSFET with metal gate and mesa isolation. Gate stack and gate dielectric are not shown for clarity and only a quarter of the channel is shown. The inset shows half of the transistor. After V. Moroz et al., *Solid State Technol.*, July 2004.

**TABLE 2.7**
Impact of stress methods on 45-nm devices. Computation is based on piezoelectric effect with 1 GPa tensile stress. After V. Moroz et al., *Solid State Technol.*, July 2004.

| Device type | S/D $Si_{0.85}Ge_{0.15}$ | STI | Cap layer | Spacer | Metal gate |
|---|---|---|---|---|---|
| n-MOS | -45% | 14% | 33% | 9% | -24% |
| n-MOS | 50% | -27% | -38% | -4% | 214% |

results. Based on data in Table 2.7, it is observed that SiGe S/D may be used to improve pMOS performance and a tensile cap layer to improve nMOS performance. The other methods can be added to these two as complementary to further enhance device performance. A uniform compressive stress of $\sim 1$ GPa is generated in the channel. Such a stress pattern is beneficial to pMOS as was successfully demonstrated in [18], but would be detrimental to nMOS. If this were the only component contributing to stress in the channel, it would have increased pMOS drain current by 70%.

The residual stresses present in thin films after deposition can be classified into two parts: (a) thermal mismatch stress and (b) intrinsic stress. Residual stresses will cause device failure due to instability and buckling if the deposition process is not controlled properly. Thermal mismatch stress occurs when two materials with different coefficients of thermal expansion are heated and expand/contract at different rates. During thermal processing, thin film materials like polysilicon, $SiO_2$, and silicon nitride expand and contract at different rates compared to the silicon substrate according to their thermal expansion coefficients.

## 2.6 Summary

In this chapter, the sources of stress and strain developed during semiconductor manufacturing are discussed. This is followed by different types of stress and strain, the state-of-the-art stress generation techniques, for both wafer-based biaxial and process-induced uniaxial stresses are discussed. Approaches to enhancing mobility by including mechanical strain in the channel layer to enhance carrier mobility and saturation velocity, and employing alternative channel materials such as SiGe and germanium with electron and hole mobilities and carrier velocities higher than those in silicon are addressed. A judicious choice of crystal orientation and current transport direction may also provide transport enhancement.

A multidisciplinary effort is required for working on novel CMOS structures

in the fields of (a) growth and fabrication of novel substrate materials, (b) device fabrication, (c) characterization, and (d) modeling. It will thus offer a complete perspective from fundamental characteristics of the new materials to device processing and performance.

## References

[1] S. Thompson, N. Anand, M. Armstrong, C. Auth, B. Arcot, M. Alavi, P. Bai, J. Bielefeld, R. Bigwood, J. Brandenburg, M. Buehler, S. Cea, V. Chikarmane, C. Choi, R. Frankovic, T. Ghani, G. Glass, W. Han, T. Hoffmann, M. Hussein, P. Jacob, A. Jain, C. Jan, S. Joshi, C. Kenyon, J. Klaus, S. Klopcic, J. Luce, Z. Ma, B. Mcintyre, K. Mistry, A. Murthy, P. Nguyen, H. Pearson, T. Sandford, R. Schweinfurth, R. Shaheed, S. Sivakumar, M. Taylor, B. Tufts, C. Wallace, P. Wang, C. Weber, and M. Bohr, "A 90 nm logic technology featuring 50 nm strained silicon channel transistors, 7 layers of Cu interconnects, low-k ILD, and 1 $\mu m^2$ SRAM cell," *IEEE IEDM Tech. Dig.*, pp. 61–64, 2002.

[2] K. Rim, S. Koester, M. Hargrove, J. Chu, P. M. Mooney, J. Ott, T. Kanarsky, P. Ronsheim, M. Ieong, A. Grill, and H.-S. P. Wong, "Strained Si NMOSFETs for high performance CMOS technology," *VLSI Technology*, pp. 59–60, 2001.

[3] C. K. Maiti, L. K. Bera, and S. Chattopadhyay, "Strained-Si heterostructure field effect transistors," *Semicond. Sci. Technol.*, vol. 13, pp. 1225–1246, 1998.

[4] A. Shimizu, K. Hachimine, N. Ohki, H. Ohta, M. Koguchi, Y. Nonaka, H. Sato, and F. Ootsuka, "Local mechanical-stress control (LMC): a new technique for CMOS-performance enhancement," *IEEE IEDM Tech. Dig.*, pp. 433–436, 2001.

[5] S. Maikap, C.-Y. Yu, S.-R. Jan, M. H. Lee, and C. W. Liu, "Mechanically strained strained-Si NMOSFETs," *IEEE Electron Dev. Lett.*, vol. 25, pp. 40–42, 2004.

[6] A. Lochtefeld and D. A. Antoniadis, "Investigating the relationship between electron mobility and velocity in deeply scaled NMOS via mechanical stress," *IEEE Electron Dev. Lett.*, vol. 22, pp. 591–593, 2001.

[7] C. S. Smith, "Piezoresistance effect in germanium and silicon," *Phys. Rev.*, vol. 94, pp. 42–49, 1954.

[8] D. Colman, R. T. Bate, and J. P. Mize, "Mobility anisotropy and piezoresistance in silicon p-type inversion layers," *J. Appl. Phys.*, vol. 39, pp. 1923–1931, 1968.

[9] G. Dorda, I. Eisele, and H. Gesch, "Many-valley interactions in n-type silicon inversion layers," *Phys. Rev. B (Solid State)*, pp. 1785–1798, 1978.

[10] S. Tiwari, M. Fischetti, P. Mooney, and J. Welser, "Hole mobility improvement in silicon-on-insulator and bulk silicon transistors using local strain," *IEEE IEDM Tech. Dig.*, pp. 939–941, 1997.

[11] P. Smeys, P. B. Griffin, Z. U. Rek, I. De Wolf, and K. C. Saraswat, "Influence of process-induced stress on device characteristics and its impact on scaled device performance," *IEEE Trans. Electron Dev.*, vol. 46, pp. 1245–1251, 1999.

[12] J. Welser, J. L. Hoyt, and J. F. Gibbons, "Electron mobility enhencement in strained-Si N-type metal-oxide-semiconductor field-effect transistors," *IEEE Electron Dev. Lett.*, vol. 15, pp. 100–102, 1994.

[13] D. K. Nayak, J. C. S. Woo, J. S. Park, K. L. Wang, and K. P. MacWilliams, "High-mobility p-channel metal-oxide-semiconductor field-effect transistor on strained-Si," *Appl. Phys. Lett.*, vol. 62, pp. 2853–2855, 1993.

[14] J. L. Hoyt, H. M. Nayfeh, S. Eguchi, I. Aberg, G. Xia, T. Drake, E. A. Fitzgerald, and D. A. Antoniadis, "Strained silicon MOSFET technology," *IEEE IEDM Tech. Dig.*, pp. 23–26, 2002.

[15] R. W. Keyes, "Explaining strain [in silicon]," *IEEE Circuits & Devices Mag.*, vol. 35, pp. 36–39, 2002.

[16] J. Welser, J. L. Hoyt, S. Takagi, and J. F. Gibbons, "Strain dependence of the performance enhancement in strained-Si n-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 373–376, 1994.

[17] V. Chan, R. Rengarajan, N. Rovedo, W. Jin, T. Hook, P. Nguyen, J. Chen, E. Nowak, X.-D. Chen, D. Lea, A. Chakravarti, V. Ku, S. Yang, A. Steegen, C. Baiocco, P. Shafer, H. Ng, S.-F. Huang, and C. Wann, "High speed 45 nm gate length CMOSFETs integrated into a 90 nm bulk technology incorporating strain engineering," *IEEE IEDM Tech. Dig.*, pp. 77–80, 2003.

[18] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson, and M. Bohr, "A 90 nm high volume manufacturing logic technology featuring novel 45 nm gate length strained silicon CMOS," *IEEE IEDM Tech. Dig.*, pp. 978–980, 2003.

[19] P. R. Chidambaram, B. A. Smith, L. H. Hall, H. Bu, S. Chakravarthi, Y. Kim, A. V. Samoilov, A. T. Kim, P. J. Jones, R. B. Irwin, M. J. Kim, A. L. P. Rotondaro, C. F. Machala, and D. T. Grider, "35% drive current improvement from recessed-SiGe drain extensions on 37 nm gate length PMOS," *IEEE VLSI Tech. Symp. Dig.*, pp. 48–49, 2004.

[20] K. Mistry, M. Armstrong, C. Auth, S. Cea, T. Coan, T. Ghani, T. Hoffmann, A. Murthy, J. Sandford, R. Shaheed, K. Zawadzki, K. Zhang, S. Thompson, and M. Bohr, "Delaying forever: uniaxial strained silicon transistors in a 90 nm CMOS technology," *IEEE VLSI Tech. Symp. Dig.*, pp. 50–51, 2004.

[21] S. E. Thompson, M. Armstrong, C. Auth, S. Cea, R. Chau, G. Glass, T. Hoffman, J. Klaus, Z. Ma, B. Mcintyre, A. Murthy, B. Obradovic, L. Shifren, S. Sivakumar, S. Tyagi, T. Ghani, K. Mistry, M. Bohr, and Y. El-Mansy, "A logic nanotechnology featuring strained-silicon," *IEEE Electron Dev. Lett.*, vol. 25, pp. 191–193, 2004.

[22] C.-H. Ge, C.-C. Lin, C.-H. Ko, C.-C. Huang, Y.-C. Huang, B.-W. Chan, B.-C. Perng, C.-C. Sheu, P.-Y. Tsai, L.-G. Yao, C.-L. Wu, T.-L. Lee, C.-J. Chen, C.-T. Wang, S.-C. Lin, Y.-C. Yeo, and C. Hu, "Process-strained Si (PSS) CMOS technology featuring 3D strain engineering," *IEEE IEDM Tech. Dig.*, pp. 73–76, 2003.

[23] K.-W. Su, Y.-M. Sheu, C.-K. Lin, S.-J. Yang, W.-J. Liang, X. Xi, C.-S. Chiang, J.-K. Her, Y.-T. Chia, C. H. Diaz, and C. Hu, "A scaleable model for STI mechanical stress effect on layout dependence of MOS electrical characteristics," *Proc. IEEE CICC Conf.*, pp. 245–248, 2003.

[24] C. Gallon, G. Reimbold, G. Ghibaudo, R. A. Bianchi, R. Gwoziecki, S. Orain, E. Robilliart, C. Raynaud, and H. Dansas, "Electrical analysis of mechanical stress induced by STI in short MOSFETs using externally applied stress," *IEEE Trans. Electron Dev.*, vol. 51, pp. 1254–1261, 2004.

[25] D. Chidambarrao, X. H. Liu, and K. W. Schwarz, "Combined dislocation and process modeling for local oxidation of silicon structure," *J. Appl. Phys.*, vol. 92, pp. 6278–6286, 2002.

[26] A. Steegen, I. De Wolf, and K. Maex, "Characterization of the local mechanical stress induced during the Ti and Co/Ti salicidation in sub-0.25 $\mu$m technologies," *J. Appl. Phys.*, vol. 86, pp. 4290–4297, 1999.

[27] A. Steegen, M. Stucchi, A. Lauwers, and K. Maex, "Silicide induced pattern density and orientation dependent transconductance in MOS transistors," *IEEE IEDM Tech. Dig.*, pp. 497–500, 1997.

[28] S. Pidin, T. Mori, R. Nakamura, T. Saiki, R. Tanabe, S. Satoh, M. Kase, K. Hashimoto, and T. Sugii, "MOSFET current drive optimization using silicon nitride capping layer for 65 nm technology node," *IEEE VLSI Tech. Symp. Dig.*, pp. 54–55, 2004.

[29] M. D. Giles, M. Armstrong, C. Auth, S. M. Cea, T. Ghani, T. Hoffmann, R. Kotlyar, P. Matagne, K. Mistry, R. Nagisetty, B. Obradovic, R. Shaheed, L. Shifren, M. Stettler, S. Tyagi, X. Wang, C. Weber, and K. Zawadzki, "Understanding stress enhanced performance in Intel 90 nm CMOS technology," *IEEE VLSI Tech. Symp. Dig.*, pp. 118–119, 2004.

[30] A. Steegen and K. Maex, "Silicide-induced stress in Si: origin and consequences for MOS technologies," *Mater. Sci. Eng.*, vol. R38, pp. 1–53, 2002.

[31] S. M. Hu, "Film-edge-induced stress in substrates," *J. Appl. Phys.*, vol. 50, pp. 4661–4666, 1979.

[32] S. L. Hsia, T. Y. Tan, P. Smith, and G. E. McGuire, "Resistance and structural stabilities of epitaxial $CoSi_2$ films on (001) Si substrates," *J. Appl. Phys.*, vol. 72, pp. 1864–1873, 1992.

[33] K. Ota, K. Sugihara, H. Sayama, T. Uchida, H. Oda, T. Eimori, H. Morimoto, and Y. Inoue, "Novel locally strained channel technique for high performance 55 nm CMOS," *IEEE IEDM Tech. Dig.*, pp. 27–30, 2002.

[34] S. L. Toh, K. P. Loh, C. B. Boothroyd, K. Li, C. H. Ang, E. Er, and L. Chan, "Reduction of local mechanical stress in a transistor using $Si_3N_4/SiO_xN_y$ contact ESL," *Electrochem. Solid-State Lett.*, vol. 8, pp. G38–G40, 2005.

[35] S. Pidin, T. Mori, K. Inoue, S. Fukuta, N. Itoh, E. Mutoh, K. Ohkoshi, R. Nakamura, K. Kobayashi, K. Kawamura, T. Saiki, S. Fukuyama, S. Satoh, M. Kase, and K. Hashimoto, "A novel strain enhanced CMOS architecture using selectively deposited high tensile and high compressive silicon nitride films," *IEEE IEDM Tech. Dig.*, pp. 213–216, 2004.

[36] A.V-Y. Thean, L. Prabhu, V. Vartanian, M. Ramon, B-Y. Nguyen, T. White, H. Collard, Q-H. Xie, S. Murphy, J. Cheek, S. Venkatesan, J. Mogab, C.H. Chang, Y.H. Chiu, H. C. Tuan, Y. C. See, M. S. Liang, and Y. C. Sun, "Uniaxial-biaxial stress hybridization for super-critical strained-Si directly on insulator (SC-SSOI) PMOS with different channel orientations," *IEEE IEDM Tech. Dig.*, pp. 509–512, 2005.

[37] R. Khamankar, H. Bu, C. Bowen, S. Chakravarthi, P. R. Chidambaram, M. Bevan, A. Krishnan, H. Niimi, B. Smith, I. Blatchford, B. Homung, J. P. Lu, P. Nicollian, B. Kirkpatrick, D. Miles, M. Hewson, D. Farber, L. Hall, H. Al Shareef, A. Varghcse, A. Gurba, V. Ukraintsev, B. Rathsack, J. DeLoach, I. Tran, C. Kaneshigc, M. Somcrvell, S. Aur, C. Machala, and T. Grider, "An enhanced 90 nm high performance technology with strong performance improvements from stress and mobility increase through simple process changes," *IEEE VLSI Tech. Symp. Dig.*, pp. 162–163, 2004.

[38] C.-H. Chen, T. L. Lee, T. H. Hou, C. L. Chen, C. C. Chen, J. W. Hsu, K. L. Cheng, Y. H. Chiu, H. J. Tao, Y. Jin, C. H. Diaz, S. C. Chen, and M.-S. Liang, "Stress memorization technique (SMT) by selectively strained-nitride capping for sub-65 nm high-performance strained-Si device application," *IEEE VLSI Tech. Symp. Dig.*, pp. 56–57, 2004.

[39] F. Nouri, P. Verheyen, L. Washington, V. Moroz, I. De Wolf, M. Kawaguchi, S. Biesemans, R. Schreutelkamp, Y. Kim, M. Shen, X. Xu, R. Rooyackers, M. Jurczak, G. Eneman, K. De Meyer, L. Smith, D. Pramanik, H. Forstner, S. Thirupapuliyur, and G. S. Higashi, "A systematic study of trade-offs in engineering a locally strained pMOSFET," *IEEE IEDM Tech. Dig.*, pp. 1055–1058, 2004.

[40] G. Eneman, P. Verheyen, R. Rooyackers, F. Nouri, L. Washington, R. Degraeve, B. Kaczer, V. Moroz, A. De Keersgieter, R. Schreutelkamp, M. Kawaguchi, Y. Kim, A. Samoilov, L. Smith, P. P. Absil, K. De Meyer, M. Jurczak, and S. Biesemans, "Layout impact on the performance of a locally strained PMOSFET," *IEEE VLSI Tech. Symp. Dig.*, pp. 22–23, 2005.

[41] F. Ootsuka, S. Wakahara, K. Ichinose, A. Honzawa, S. Wada, H. Sato, T. Ando, H. Ohta, K. Watanabe, and T. Onai, "A highly dense, high-performance 130 nm node CMOS technology for large scale system-on-a-chip applications," *IEEE IEDM Tech. Dig.*, pp. 575–578, 2000.

[42] S. Ito, H. Namba, K. Yamaguchi, T. Hirata, K. Ando, S. Koyama, S. Kuroki, N. Ikezawa, T. Suzuki, T. Saitoh, and T. Horiuchi, "Mechanical stress effect of etch-stop nitride and its impact on deep submicron transistor design," *IEEE IEDM Tech. Dig.*, pp. 247–250, 2000.

[43] L. K. Bera, W. Y. Loh, L. H. Guo, X. W. Zhang, G. Q. Lo, N. Balasubramanian, and D.-L. Kwong, "The impact of uniform strain applied via bonding onto plastic substrate on MOSFET performance," *IEEE Electron Dev. Lett.*, vol. 27, pp. 58–61, 2006.

[44] K. Uchida, R. Zednik, C.-H. Lu, H. Jagannathan, J. McVittie, P. C. McIntyre, and Y. Nishi, "Experimental study of biaxial and uniaxial strain effects on carrier mobility in bulk and ultrathin-body SOI MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 229–232, 2004.

[45] K. Uchida, H. Watanabe, A. Kinoshita, J. Koga, T. Numata, and S. Takagi, "Experimental study of carrier transport mechanism in ultrathin-body SOI n- and p-MOSFETs with SOI thickness less than 5 nm," *IEEE IEDM Tech. Dig.*, pp. 47–50, 2002.

[46] J. P. Peng, D. Chidambarrao, and G. R. Srinivasan, "Novel: a nonlinear viscoelastic model for thermal oxidation of silicon," *COMPEL*, vol. 10, pp. 341–353, 1991.

[47] V. Senez, D. Collard, and B. Baccus, "Analysis and application of a viscoelastic model for silicon oxidation," *J. Appl. Phys.*, vol. 76, pp. 3285–3296, 1994.

[48] H. Umimoto and S. Odanaka, "Three-dimensional numerical simulation of local oxidation of silicon," *IEEE Trans. Electron Dev.*, vol. 38, pp. 505–511, 1991.

[49] S. Cea and M. E. Law, "Three dimensional nonlinear viscoelastic oxidation modeling," *Proc. SISPAD*, pp. 97–98, 1996.

[50] H. Rueda, S. Cea, and M. E. Law, "Mechanical stress modeling for silicon fabrication processes," *Proc. SISPAD*, pp. 53–55, 1997.

[51] G. Scott, J. Lutze, M. Rubin, F. Nouri, and M. Manley, "NMOS drive current reduction caused by transistor layout and trench isolation induced stress," *IEEE IEDM Tech. Dig.*, pp. 827–830, 1999.

[52] M. E. Law and S. M. Cea, "Continuum based modeling of silicon integrated circuit processing: an object oriented approach," *Comp. Mat. Sci.*, vol. 12, pp. 289–308, 1998.

[53] S. M. Cea, M. Armstrong, C. Auth, T. Ghani, M. D. Giles, T. Hoffmann, R. Kotlyar, P. Matagne, K. Mistry, R. Nagisetty, B. Obradovic, R. Shaheed, L. Shifren, M. Stettler, S. Tyagi, X. Wang, C. Weber, and K. Zawadzki, "Front end stress modeling for advanced logic technologies," *IEEE IEDM Tech. Dig.*, pp. 963–966, 2004.

[54] V. Moroz, N. Strecker, X. Xu, L. Smith, and I. Bork, "Modeling the impact of stress on silicon processes and devices," *Mat. Sci. Semicond. Proc.*, vol. 6, pp. 27–36, 2003.

[55] Z. Krivokapic, V. Moroz, W. Maszara, and M.-R. Lin, "Locally strained ultrathin channel 25 nm narrow FDSOI devices with metal gate and mesa isolation," *IEEE IEDM Tech. Dig.*, pp. 445–448, 2003.

# 3

## *Strain-Engineered Substrates*

An engineered substrate is a material that can be fabricated and introduced in the conventional silicon manufacturing, resulting in products that are unique and could not have been fabricated using only silicon substrate. Engineered substrates generally shall have different lattice constants. In 1992, several researchers demonstrated that controlled strain could result in up to a doubling of speed in manufacturable CMOS transistors. Mobility enhancement in Si has been demonstrated utilizing engineered substrates, for example, by straining the Si channel. Until recently, researchers have focused on biaxial strain, created by growing high-quality layers of strained material across the entire wafer. Because of the conflicting effects that strain has on n- and p-MOSFETs, the whole-wafer approach of building strain into silicon typically improves the performance of one transistor type more than that of the other, which has held back the commercial use of biaxial strain. However, electron and hole mobility can be optimized separately by patterning substrates to expose different Si crystal faces for each device.

Another approach toward strain engineering has been to selectively strain parts of individual transistors. Using this local effect, known as uniaxial strain engineering, one can impose either tensile or compressive strain on both n- and p-MOSFETs on the same wafer and hence the CMOS circuitry performance gets better as a whole. It is expected that by combining the biaxial strain with uniaxial strain will enable one to compensate for the uneven effect that substrate strain has on the different types of CMOS devices. It is important to note that biaxial strain offers some advantages, such as uniformity of strain for all transistors, which uniaxial strain cannot provide.

In this chapter, we discuss the present trends and applications of strained-Si films in SiGe-based CMOS technology. In-depth discussion will cover the epitaxial (both homo- and heteroepitaxy) film growth and characterization of the strained-Si layers on virtual substrates. The status of the use of several high-mobility substrates and their fabrication will be covered. The impact of defects on the electrical device performance such as leakage current and carrier lifetime will be addressed. Principle as well as the corresponding engineered substrates and their fabrication process will be presented.

A large number of characterization is necessary for strained layers, i.e., strain state, thickness uniformity, defectivity, and mobility data will be discussed. We will focus on SSOI, giving a technical status on the wafer quality presenting a wide set of data obtained through specific characterization tech-

niques, including crystal defect revelation on thin films.

## 3.1 Epitaxy

For the past 40 years, the material of choice for the semiconductor industry has been Si. Currently, almost all semiconductor-based systems are built on a particular bulk substrate material with constant lattice. Besides creating new materials on a silicon substrate, new lattices of the bulk lattice constants can be used to strain thin layers of traditional materials, such as silicon and germanium. The limitations of lattice-matched semiconductor systems are the nonaccessability of the rich nature of semiconductor bandgaps and lattice constants. Lattice engineering allows one to engineer bandgap and structures and can be used to strain thin layers of traditional materials such as silicon and germanium. In general, electronic performance improves as the lattice constant increases within a class of semiconductor materials. For example, Si is used for digital applications, whereas III-V and II-VI materials are used mostly for high frequency and optical applications. The lattice constants inbetween Si and Ge offer a host of new materials and devices that can be constructed in the Si CMOS manufacturing infrastructure. These performance limitations forced the evolution of lattice-mismatched systems which will allow the integration of any material (film) on any bulk semiconductor substrate. In early epitaxial growth research, many lattice-mismatched films were deposited on bulk semiconductors, but the lattice-mismatch between the film and substrate led to poor material quality. Thus, nearly all electronic and optoelectronic systems are still being built on fixed lattice constant bulk semiconductors.

During the last 20 years, various process options that were being considered include: strain engineering - both global (wafer level) and local (process strain) and different engineered substrate options where thin silicon layers are transferred onto silicon substrates. Staying within the constraints of lattice-matching, researchers in the mid-1970s began to explore incorporating slightly mismatched semiconductor films by keeping the level of mismatch and the film thickness below the critical thickness. In particular, the most important engineered substrate now is the SiGe-Si system which has migrated into applications for the strained-SiGe base heterojunction bipolar transistors (HBTs). Parallel research efforts over the past 20 years in epitaxy, strain-relaxed semiconductor epitaxy, and wafer-bonding technology are ushering in a new age of lattice-mismatched substrate engineering. Some examples are SOI wafers, SiGe, and strained-Si. With the flexibility to combine thin highly strained layers and films composed of previously unattainable lattice constants, one can now fabricate materials that have not existed before that may find applications in microelectronics.

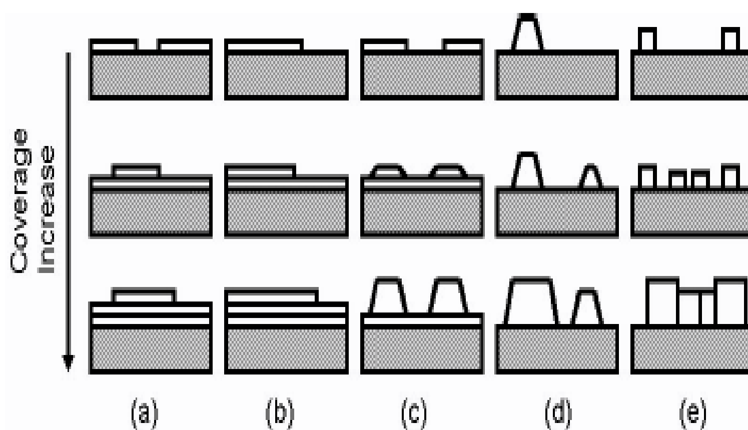



**FIGURE 3.1**
Schematic representation of the five crystal growth modes most frequently occurring on flat surfaces of substrate crystals. (a) Layer-by-layer or Frank-van der Merwe (FM mode); (b) step flow (SF mode); (c) layer plus island or Stranski-Krastanov (SK mode); (d) island or Volmer-Weber (VW mode); and (e) columnar growth mode (CG mode). After O. L. Guise, *Production and Characterization of Novel Nanostructure Materials*, Ph.D. Thesis, University of Pittsburgh, 2005.

The concept of epitaxy was introduced by L. Royer in 1928 to describe the growth of crystalline nuclei on various solid substrates. Epitaxy, in general, means the ordered growth of a crystalline or, at least, highly oriented material over another one, acting as substrate. The increase of a solid by addition of new atoms or molecules of the same material on top of its surface is called "homoepitaxy." Oriented or single crystalline growth of layer occurs on a substrate of a different material is known as "heteroepitaxy."

Heteroepitaxy is the process of growing a film with equilibrium lattice constant that is different from that of the substrate. Heteroepitaxy of semiconductor materials has been an active area of research for the last four decades. The interest is driven by the possibility of creating novel heterostructure electronic and optical devices [1], as well as integrating existing devices in different material systems, leading to the production of integrated circuits with an increased functionality and at a lower cost. The foundation of heteroepitaxy was laid by two important contributions. Frank and van der Merwe in 1949 [2] showed theoretically that if a lattice mismatched layer is grown on a thick substrate, the layer will be pseudomorphic provided the mismatch is small and thickness of the layer is not large. Shockley suggested the use of semiconductors of different bandgaps for fabrication of heterostructure devices [3].

Heteroepitaxy is widely employed in semiconductor technology, through the

choice of proper materials and alloys, artificial structures, such as quantum wells and superlattices, can be grown with precisely tuned properties. In the epitaxial approach, when an alloy with a different lattice constant is epitaxially grown on substrate, initially, the growth is coherent, however, dislocations, called misfit dislocation, form with increasing thickness. As the layer thickness of the epitaxial film is increased, the total strain energy of the system increases and, at some critical thickness, it would be energetically favorable to relax the lattice by dislocation formation. Understanding the nucleation and propagation of dislocations introduced into lattice mismatched semiconductors is important in developing techniques for creating lattice engineered substrates. How misfit dislocations are introduced, i.e., the kinetics of dislocation introduction, is an active area of research, especially experimentally measuring the velocities of dislocations in strained layers. In addition to the velocity of dislocations, nucleation plays an important role, specifically the presence or absence of heterogeneous nucleation sources.

The main factors contributing to the appearance of a growth mode over another one are the lattice mismatch between material A and material B, the growth temperature, the evaporation rate and the surface energies of both materials as well as the interface energy. Heteroepitaxial films can grow in the following modes, as shown schematically in Fig. 3.1. One can easily distinguish five growth modes (most frequently occurring) upon deposition of a material B on a substrate A. These growth modes are: (a) Frank-van der Merwe (FM) or layer by layer; (b) step-flow mode (SF); (c) Stranski-Krastanov (SK) or layer plus island; (d) Volmer-Weber (VW) or island growth mode; (e) columnar growth mode (CG). Another type of heteroepitaxy, named van der Waals (vdW) epitaxy occurs when the film and substrate interact only through van der Waals forces such as growth of organic thin films on inorganic substrates.

In heteroepitaxy, FM, SK, and VW growth modes are the most important. The Frank-van der Merwe (FM), or layer-by-layer mode, leads to smooth surfaces. It is characterized by the fact that each monolayer starts growing by 2D islands, after the previous one is completed. FM growth is favored when the misfit is very low and when the surface free energy of the film is lower than the sum of the one of the substrate and of the interface free energy (full wettability). In fact, the growth mode also depends on the elastic strain $\epsilon$ in the film which depends on the lattice mismatch between the substrate A (lattice constant $a_A$) and the film B (lattice constant $a_B$): $\epsilon = (a_B - a_A)/a_A$. The lattice mismatch between A and B should be small for the FM growth mode to occur. It is very common in homoepitaxy and is often observed when metals grow on metal substrates.

In Frank-van der Merwe (FM) or layer-by-layer growth mode will occur if, upon deposition of a material B on a substrate A, the atoms or molecules of B bound more strongly to A than to each other. After the formation of a monolayer, the substrate is covered with a second monolayer which is less tightly bound than the first one. From a simplified energetic standpoint, if

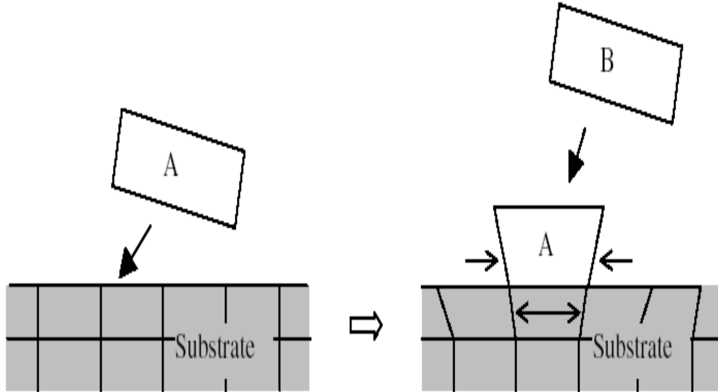



**FIGURE 3.2**
Schematic diagram of epitaxial strain relaxation via island formation.

the surface free energy of B is lower than the surface free energy of A, the interface energy between A and B being neglected, then the material B will wet the substrate A and layer-by-layer growth occurs.

The Volmer-Weber (VW), or island mode is the opposite one, which leads to the formation of 3D islands nucleating over the substrate surface. Volmer-Weber or Island Growth Mode involves the direct nucleation of 3D islands or clusters of material B on substrate A. The driving mechanism is opposite to the one for the FM mode: the atoms or molecules of B bound stronger to each other than to A. Again, from a simplified energetic standpoint, the surface free energy of B is larger than the surface free energy of A, the interface energy between A and B being neglected. Again, the growth mode also depends to some extent on the elastic strain $\epsilon$ in the film. In terms of surface free energy, VW mode occurs when the surface energy of the film is larger than the ones of the substrate and interface, as it is frequently the case when the elastic energy produced by a sizable misfit is included. This mode is often observed for the growth of metals on insulators, including metals on alkali halides and graphite.

An intermediate situation is the Stranski-Krastanov (SK), or layer-plus-island mode, and it is characterized by the layer-by-layer growth of a few monolayer (the so-called wetting layer), followed by the nucleation of 3D crystalline islands on top of them. Because of the lattice mismatch between substrate A and deposited material B, the strain accumulated in the wetting layer of B on A linearly increases with the thickness of the wetting film. After the formation of a few wetting layers of material B on substrate A, the subsequent layer growth becomes unfavorable and the growth mode converts to the

formation of islands (see Fig. 3.2). It is important to have a balance between the energy that would be gained from relaxing with 3D islands and the free surface energy building up. The SK mode is the most widespread growth mode in heteroepitaxy. This layer-plus-island process is the most frequent one: It is produced by an increasing imbalance with thickness between the tendency to wet the substrate surface and the accumulation of elastic energy by misfit in the film. Strain plays an important role in SK mode. The energy responsible for the growth of islands increases considerably with strain. The driving mechanism is the very efficient strain energy relaxation possible with the formation of 3D islands by lateral expansion or compression in the direction of the free faces because the material B conforms to the lattice of the substrate A for the first layers.

The growth of SiGe on Si(100) proceeds via the SK growth mode. In the early stages of Ge deposition on Si(100), i.e., below the 2D/3D threshold, the formation of a Ge wetting layer takes place. After the formation of a Ge wetting film, whose thickness depends on the experimental conditions (substrate temperature and deposition rate), 3D Ge islands nucleate randomly on the surface by self-assembly to relieve the strain accumulated in the Ge wetting layer. Growth kinetics play an important role in the 2D-3D transition so the substrate temperature can be used to tune the morphology of the resulting islands. One can distinguish the three regimes: below transition, after transition and formation of islands, and coalescence of islands. The difference in growth mode between SiGe islands (Stranski-Krastanov) and SiGeC islands (Volmer-Weber) can be easily characterized by RHEED, x-ray scattering or cross-sectional TEM.

Ge is deposited on a Si surface modified by misfit dislocations generated in buried compositionally graded $Si_{1-x}Ge_x$ layers. Two types of such layers are typically made. The first type consists of linear graded films where the Ge composition is increased with a grading rate of 10% per micrometer. A Si layer caps the graded structure. The dislocation strain accumulated in the buried system propagates through the Si capping layer. Fitzgerald et al. [4] have shown that a nonuniform distribution of misfit dislocations in heteroepitaxial systems generates a nonuniform distribution of elastic strain at the surface.

## 3.2 Heteroepitaxy and Strain Control

An important parameter in heteroepitaxy is the geometric misfit f, expressing (usually in percentage) the amount of mismatch between the two structures, along one surface vector or the other one:

$$f = \frac{a_B - a_A}{a_A} \tag{3.1}$$

where $a_A$ and $a_B$ are the unstrained lattice parameters of the two layers. Germanium has a 4.2% larger lattice constant than silicon. The lattice constant of relaxed- or bulk-$Si_{1-x}Ge_x$ films ($0 < x < 1$) is almost linear as predicted by Vegard's law although small variations to this have been measured. An approximation accurate to about $10^{-4}$ nm is given by [5]:

$$a_{Si_{1-x}Ge_x} = 0.5431 + 0.01992x + 0.0002733x^2 \quad nm. \tag{3.2}$$

If a thin $Si_{1-x}Ge_x$ film is grown on top of a $Si_{1-y}Ge_y$ film then for $x > y$ the top layer is compressively strained (Fig. 3.3(b)) while for $x < y$ the layer is tensile strained (Fig. 3.3(d)). In both cases the strain is in the plane of the layer ($\epsilon_x = \epsilon_y = \epsilon_\parallel$) but this strain also produces a perpendicular strain, $\epsilon_\perp$, resulting in a tetragonal distortion to the lattice (Fig. 3.3(b)). In isotropic elasticity theory the strains are related by Poisson ratio, $\nu$, through

$$\epsilon_\perp = \frac{-2\nu}{1-\nu}\epsilon_\parallel \tag{3.3}$$

If the unstrained lattice parameters of the two layers are defined as $a_A$ and $a_B$, respectively, with thicknesses $h_A$ and $h_B$ and assuming $a_A < a_B$ (Fig. 3.3 (a)) then the tetragonal distortion produces a parallel lattice constant (Fig. 3.3(b)),

$$a_\parallel = a_A\left[1 + \frac{f}{\left(1 + \frac{G_A h_A}{G_B h_B}\right)}\right] \tag{3.4}$$

where $G_i$ are the shear moduli of each layer and $h_i$ are the thicknesses. The in-plane strain in layer A is given by the term and so in equilibrium the in-plane strain of layer A, $\epsilon_\parallel^A$, is related to the strain in layer B, $\epsilon_\parallel^B$ through

$$\epsilon_\parallel^A = -\left(\frac{G_B h_B}{G_A h_A}\right)\epsilon_\parallel^B \tag{3.5}$$

Since the stress imposed by the substrate, $\sigma_\parallel$, is 2D and the component perpendicular to the surface, $\sigma_\perp$, is zero, the epitaxial film is free to rearrange its lattice parameter perpendicular to the surface, in such a way that its internal energy is minimized. The resulting strain components are $\epsilon_\parallel$ and $\epsilon_\perp$, parallel and perpendicular to the surface, respectively. They are opposite in sign, since the natural tendency of the strained structure is to preserve the volume of the unit cell. In fact, they can be expressed in terms of stress components, via the elastic equations involving the Poisson ratio $\nu$ and the Shear modulus $\mu$, which are in turn related to the bulk modulus B, that is, the stiffness of the material.

As long as the film is strained in such a way that its in-plane lattice constants match those of the substrate, the modulus of the parallel strain equals the misfit. As the thickness h increases, the elastic energy increases correspondingly, up to a limit $h_c$ above which the uniformly strained state is no

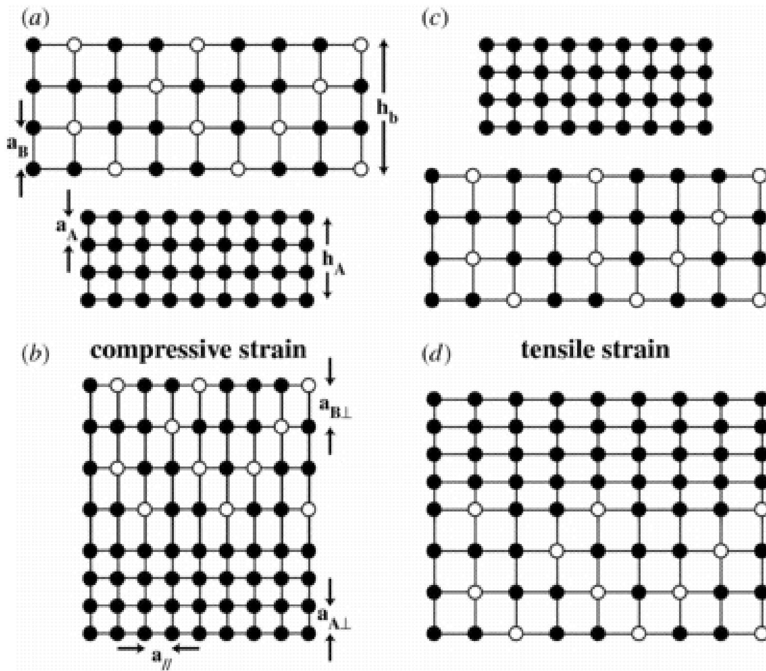



**FIGURE 3.3**
(a) A schematic diagram of the bulk lattice constant of a thin $Si_{1-x}Ge_x$ film to be grown on top of a thin bulk-silicon layer. (b) A schematic diagram showing the tetragonal lattice distortion when the two films from (a) are placed together with the top $Si_{1-x}Ge_x$ film being compressively strained. (c) A schematic diagram of the bulk lattice constants of a bulk-Si film to be grown on top of a bulk-$Si_{1-y}Ge_y$ film. (d) A schematic diagram of the two films in (c) placed together with the top film being tensile strained. After D. J. Paul, *Semicond. Sci. Technol.*, Vol. 19, 2004(R75-R108).

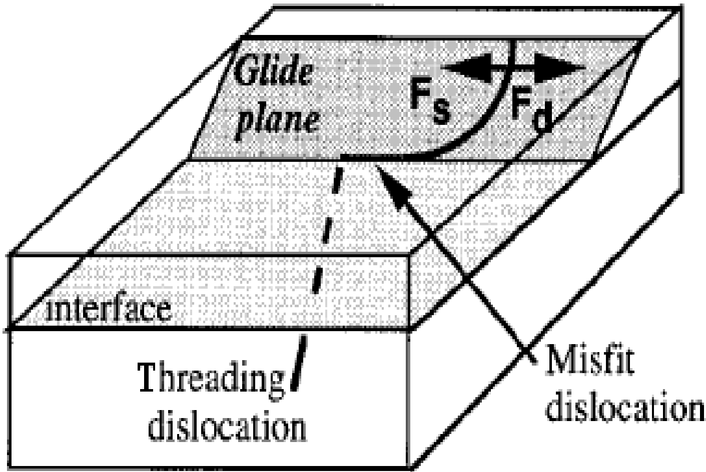



**FIGURE 3.4**
Threading dislocation as source of misfit dislocation. $F_s$ and $F_d$ are, respectively, the force due to the self-stress and the Peach-Koelher forces.

longer energetically favorable. The system can lower its total free energy instead, by generating so-called misfit dislocations (Fig. 3.3), in order to relieve part of the strain.

In practice, the calculation of the critical thickness $h_c$ is not an easy task, since several contributions enter the energy balance. In the simplest approach, one assumes thermodynamic equilibrium and takes into account just two elastic energy contributions: (a) the energy due to the residual homogeneous strain $\epsilon_{\|} < f$ and (b) the energy stored in the inhomogeneous strain field of a square grid formed by two independent sets of parallel dislocation lines.

For a given layer/substrate system with a misfit f, there will be a special thickness, known as the critical thickness, $h_c$ first predicted by Frank and van der Merwe [2], at which introducing an interfacial dislocations array will reduce the energy of the system. Another and strictly equivalent approach of this problem is to consider the development of an interfacial dislocation from a pre-existing threading dislocation (Fig. 3.4) driven by the misfit stress in the layer opposing its self-stress. An implicit expression for $h_c$ at a given misfit f is obtained by setting to zero the first derivative of the total elastic energy, at zero dislocation density:

$$h_c = \frac{b(1 - \nu cos^2\theta)}{8\pi f(1+\nu)cos\lambda}\left(ln\frac{4t_c}{b}\right) \tag{3.6}$$

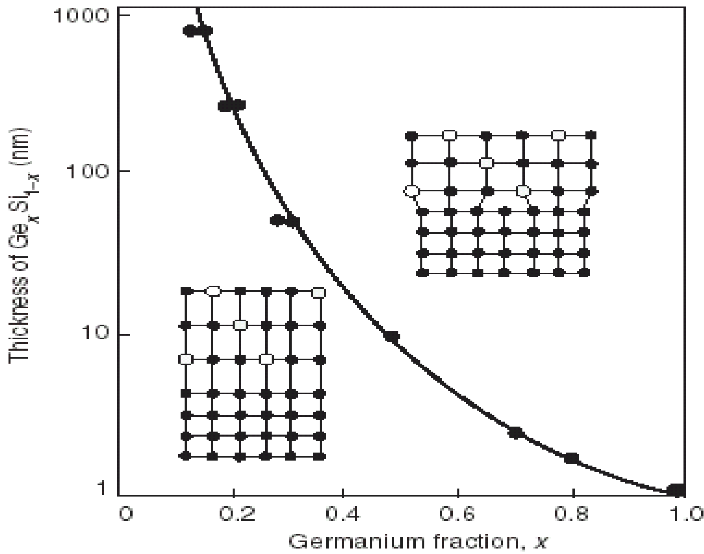



**FIGURE 3.5**
Experimental dependence of the critical thickness with the misfit, as predicted by Eqn. 3.6 for SiGe/Si(001).

where $\lambda$ is the angle between the slip direction and that direction in the interface which is perpendicular to the line of intersection of the slip plane with the interface, $\theta$ the angle between the dislocation line and its Burgers vector, b the Burgers vector modulus, and $\nu$ is the Poisson coefficient.

Examination of experimental results of heteroepitaxial growth revealed that generally the transition, between the strained unrelaxed state and the relaxed one (at least partially), occurred at larger thickness than predicted by the simple equilibrium theory [6, 7]. This discrepancy is high, if the misfit is low and also the growth temperature. This fact is obviously related to kinetic factors inhibiting the formation of the misfit dislocation arrays. Three principal kinetic models have been proposed [8, 9, 10], and are based on the existence of energy barriers involved in the formation, motion, and multiplication of misfit dislocations. All these models allow to bring the theoretical critical thickness closer to the experimental points, either in the case of low mismatched systems [8, 9] or in the case of higher misfit [10]. In Fig. 3.5, experimental dependence of the critical thickness $h_c$ with the misfit f for $Si_{1-x}Ge_x$ layers on Si(001), where the misfit is proportional to the Ge fraction, is shown. Experimentally it was observed that many pseudomorphic layers could be grown well above the critical thickness values predicted from the above equilibrium theories.

A number of models have been developed to predict the critical thickness of the strained epitaxial layer. van der Merwe [11] reported a thermodynamic

equilibrium model by minimizing the total energy of a system with the generation of a periodic array of dislocations. This produced a critical thickness defined as

$$h_c \simeq \frac{19}{16\pi^2}\left(\frac{1+\nu}{1-\nu}\right)\left(\frac{b}{f}\right) \tag{3.7}$$

where b is the slip distance which for misfit dislocations is the Burgers vector. For a bulk-silicon substrate b = 0.4 nm and more generally $a_A/\sqrt{2}$ where $a_A$ is the lattice constant of the relaxed substrate.

Matthews and Blakeslee used an equivalent approach of balancing the forces for a propagating and threading dislocation [12]. Balancing the force of the threading dislocation arm in an uncapped epilayer with the restoring line tension force from the self-energy of the extra interfacial dislocation created during relaxation and solving produces a critical thickness of

$$h_c \simeq \frac{b}{2\pi f}\frac{(1-\nu cos^2\theta)}{(1+\nu)cos\lambda}\left[ln\left(\frac{h_c}{b}+1\right]\right. \tag{3.8}$$

where $\theta$ is the angle between the dislocation line and its Burgers vector and $\lambda$ is the angle between Burgers vector and the direction in the interface normal to the dislocation line. By taking the misfit between pure Si and Ge then f = 0.0418, cos $\lambda$ = cos $\theta$ = 0.5 for $60^o$ mixed a/2<110> dislocations, Eqn. (3.8) reduces to

$$h_c \simeq \frac{0.55}{x}ln(10h_c) \quad (h_c \simeq 1.7793x^{-1.2371} \quad nm) \tag{3.9}$$

More reliable predictions were reported by Houghton [9] in an extensive study which also considered the nucleation of dislocations in single and multiple epilayers. A general expression for the extent of plastic strain relaxation, $\Delta\epsilon$(t), of a strained layer under an arbitrary thermal cycle of temperature T in time, t, was also developed, producing

$$\Delta\epsilon(t) = \frac{B\nu N_o t^2 cos\lambda}{2}\left(\frac{\tau_{eff}}{\mu}\right)^y exp\left[\frac{-(Q_N+Q_v)}{k_B T}\right] \tag{3.10}$$

where B is a material constant, v is the misfit dislocation extension velocity, $N_o$ is the density of incipient nucleation centers for dislocations at t = 0, $Q_N$ is the activation energy for nucleation of a misfit dislocation, $Q_v$ is the activation energy for a/2<110> $60^o$ type dislocation glide, $\tau_{eff}$ is the effective stress, y is a material parameter equal to 4.5, $\mu$ is the shear modulus and $k_B$ is Boltzmann's constant. $Q_n$ + $Q_v$ has been measured to be 4.75 eV and therefore the plastic strain relaxation has a large variation with temperature. Adding the experimentally extracted parameters into Eqn. (3.10) reduces the equation to

$$\Delta\epsilon(t) = 9.4\times 10^3 N_o t^2(\tau_{eff})^{4.5} exp\left[-\left(\frac{4.75}{k_B T}\right)\right] \tag{3.11}$$

**TABLE 3.1**
Strained-Si film requirements.

| | |
|---|---|
| Particle levels | <200@0.25$\mu$m |
| (limited by measurement technique) | <20@0.5$\mu$m |
| Threading Dislocation Density (TDD) | $\sim$ 10E4 |
| Pile-up Dislocation Density | 0 (cm/cm$^2$) |
| Surface Roughness | 10×10; R$_a$ 7.5 Å |
| | 1×1; R$_a$ 0.5 Å |
| Germanium Uniformity | <0.5% |
| Strained-Si Thickness Uniformity | < 2% |
| Fully Relaxed Layer | < 99% |

One significant result of this work was demonstrating that a cap layer of the same composition as the substrate does not provide any significant protection to the strained layer from relaxation on account of cancellation of the strain fields of the dislocations produced. Through analysis of data from a number of different types of epitaxially grown SiGe layers, the misfit extension velocity for a/2<110> 60$^o$ dislocations was demonstrated to be independent of the strain layer geometry, Ge fraction of the epilayer and the effective stress. An empirical equation with the velocity and $\tau_{eff}$ in MPa was determined to be

$$\nu = 9.8 \times 10^4 (\tau_{eff})^2 exp\left[-\left(\frac{2.25}{k_B T}\right)\right] \tag{3.12}$$

This study also demonstrated that the nucleation and glide of misfit dislocations in Si$_{1-x}$Ge$_x$/Si heterostructures are independent, thermally activated processes in the initial stages of strain relief when the nucleations center density, N(t), is low ($\leq 10^4$ cm$^{-1}$). The nucleation rate for misfit dislocations was determined to be

$$\frac{dN(t)}{dt} = 0.7 \times 10^5 (\tau_{eff})^{2.5} exp\left[-\left(\frac{2.5}{k_B T}\right)\right] \tag{3.13}$$

If the misfit is within some percent, as in case of Ge on Si(001) with f (4.2%), elastic relaxation takes place before the critical thickness for dislocation nucleation is reached. In this case, $\epsilon_{\parallel}$ is lowered via surface roughening, which allows for local lateral expansion of the lattice parameter at the edges of 2D islands, or at the facets of 3D aggregates, including the case of surface modulations.

The requirements for engineered substrates may be summarized as follows:

• Surface quality must be appropriate for leading edge processes. Needs to be comparable to conventional Si substrate,

• Material must be compatible with silicon-based process technologies and have relatively few integration issues,

• Must also be compatible with other materials systems such as SOI to address future technology nodes,

- Price must be competitive, e.g., cost advantage over purchase of next-generation lithography for equivalent performance enhancement, and
- Ge fraction must be high enough to provide pMOS and nMOS enhancement while maintaining a workable strained-Si layer thickness.

## 3.3 Engineered Substrates: Technology

The huge costs of scaling CMOS devices according to Moore's law have left the silicon industry at a crossroad. In order to improve the speed of VLSI/ULSI circuits, new materials and device structures are being proposed. Recent progress in the heteroepitaxial deposition of group-IV alloys, specifically $Si_{1-x}Ge_x$ and its variants on Si, has led to the application of bandgap engineering techniques to silicon-based heterostructure bipolar and field-effect devices. Currently, the majority of the SiGe heterobipolar devices are basically a Si-BJT with Ge selectively introduced in the base region. This means that existing Si fabrication processes can be easily modified to produce SiGe-HBTs. The only additional process is depositing Ge in the base region. However, stability is a major concern for the epitaxially grown SiGe layers as one cannot exceed the critical layer thickness for a given Ge mole fraction. Exceeding this limit leads to the formation of defects that can result in an unstable or metastable film and ultimately render the device useless.

A major issue for the device designers is to achieve symmetrical electrical operation from equivalently sized n- and p-MOSFETs for increased packing density in CMOS circuits. However, the p-channel (pMOS) devices are inferior to the n-channel (nMOS) ones in terms of current drive capability and speed performance. This is a consequence of lower mobility of holes compared to that of electrons in Si. Thus, size of the pMOS device is made 2-3 times larger compared with nMOS device, thereby affecting the integration density and speed. A higher mobility in the channel of a p-MOSFET will improve both circuit speed and the level of integration in CMOS. One possibility is to fabricate Si heterostructure devices using group-IV semiconductor materials on Si substrates.

Germanium and carbon are widely used heterojunction partners with Si. The lattice mismatch between Ge and Si, and Si and C is approximately 4.2% and 52%, respectively. There is a significant band gap difference between Si (1.12 eV), Ge (0.67 eV) and C (5.5 eV) which allows the possibility of bandgap engineering in group-IV alloys. SiGe is a Si-based alloy, which has been widely studied in the last decade. strained-Si/$Si_{1-x}Ge_x$/Si heterojunctions and quantum wells have been employed in manufacturing many microelectronic and optoelectronic devices, such as heterojunction bipolar transistors, MOSFETs, infrared photodetectors, and optical waveguide devices. In

all these applications, the heterointerface between Si and SiGe plays a crucial role. The addition of C has demonstrated further modification of the electronic and optical properties of Si/SiGe alloys. Carbon varies the degree of strain in the Si/SiGe system, an important parameter in the optoelectronic properties and thermal stability of the alloy.

The fundamental limitation of SiGe alloy films on Si substrate is the lattice mismatch between the film and substrate, especially for $Si_{1-x}Ge_x$ with high Ge content. The constrain in critical layer thickness of strained $Si_{1-x}Ge_x$ layers has imposed severe limitations on device structures, restricting their application only for low Ge concentration, thin active layer and relatively low process temperature windows. An interesting approach that has been used is the addition of small amount of C in SiGe films to perform strain engineering [13]. Since C is a much smaller atom than Si, while Ge is larger, C provides dilational strain, which compensates the compressive strain provided by Ge and hence increases the critical thickness as shown in Fig. 3.5. Estimates of how much C is needed for complete strain compensation in SiGe films range from Ge-to-C ratios of 8:1 to 13:1, with a gradual consensus building up at a Ge-to-C ratio of 8.5:1 [14]. Therefore, incorporating relatively small-sized carbon atoms substitutionally into SiGe system enables one to compensate the strain, allowing more flexibility in strain and bandgap engineering.

The advances in the growth of strained-Si layers on relaxed-SiGe buffer layers, combined with higher values of both the hole and electron mobilities in strained-Si, have led to increased interest in silicon-based heterojunction field-effect transistors using conventional Si-processing technology. Strained-Si technology enables improvements in CMOS performance and functionality via replacement of the bulk, cubic-crystal Si substrate with a Si substrate that contains a tetragonally distorted, biaxially strained-Si thin film at the surface. Due to changes in its crystalline structure (i.e., its symmetry is different due to its strain state), the strained-Si film has electronic properties that are superior to those of bulk-Si. Specifically, the strained-Si film has greater electron and hole mobilities, which translate into greater drive current capabilities for nMOS and pMOS transistors, respectively. This increased transistor drive current can be traded off for a reduction of dynamic power consumption [15, 16]. The ability to tailor the performance of strained-Si technology for each specific application will have a far-reaching impact, from high-speed products like microprocessors and field-programmable gate arrays, to the low-power products used in wireless communications and other mobile devices.

The key to enabling high-performance devices is the fabrication of low-defect density strained-Si films. The lattice mismatch in SiGe material system is 4.2%, resulting in a very high misfit and threading dislocation density. Early work on the nature of misfit dislocation nucleation has revealed that high mismatch strain nucleate many threading dislocations locally, resulting in a high threading dislocation density. However, a slow introduction of strain at high temperature has shown that dislocation nucleation could be controlled (i.e.,

no rampant nucleation) and dislocation propagation could be encouraged, resulting in a more global means of strain relief. The combination results in slightly mismatched layers with a high degree of relaxation and low threading dislocation density. Most of the research has concentrated on devices having strained layers with thicknesses below the critical thickness. Growing a Si layer on a SiGe film, which has a larger lattice constant than Si, generates the strained-Si heterostructure. The challenge faced when forming this heterostructure is the engineering of the lattice constant of the SiGe film. The challenge faced when forming this heterostructure is the engineering of the lattice constant of the SiGe film.

Among the lattice-mismatched heterostructures, SiGe heterostructures are usually formed via coherent structures, in which the SiGe layer is pseudomorphic to the underlying silicon substrate being subjected to misfit strain and structures incorporating a virtual substrate (VS), designed to accommodate the lattice mismatch between the silicon substrate and the overlying active layer. When a thin film with a larger lattice constant (e.g., $Si_{1-x}Ge_x$) is grown on a smaller lattice constant substrate (e.g., silicon), the film maintains in-plane lattice constant of the substrate and is under a biaxially compressive strain. This is known as the type-I band alignment where virtually all the entire band offset occurs in the valence band with minimal band offset in the conduction band. This type of structure, favorable for hole confinement has been exploited in several novel heterostructure devices viz., buried channel p-MOSFETs, p-MODFETs and HBTs (see for example, excellent reviews by Paul [17] and Konig and Daembkes [18]).

Similarly, a smaller lattice constant silicon epilayer will be under biaxial tension when grown on a larger lattice constant relaxed-$Si_{1-x}Ge_x$ substrate. In this case, type-II band offset occurs and the structure has several advantages over the more common type-I band alignment, as a large band offset is obtained in both the conduction and valence bands, relative to the relaxed-$Si_{1-x}Ge_x$ layer [19]. This allows both electron and hole confinements, making it useful for both n- and p-type devices for strained-Si/SiGe-based CMOS technology. Since strained-Si provides both larger conduction and valence band offsets and does not suffer from alloy scattering (mobility degradation) [20]), a significant improvement in carrier mobility can be achieved. Strained-Si is more difficult to grow as compared to strained-$Si_{1-x}Ge_x$, since a $Si_{1-x}Ge_x$ substrate is not easily available and until recently, growth of relaxed-$Si_{1-x}Ge_x$ without forming a large concentration of defects due to dislocation is difficult. Studies on the incorporation of a small amount of C atoms in Si/SiGe material system to develop new types of buffer layers with reduced misfit dislocations may be useful [21].

### 3.3.1 Virtual Substrates

Jorke and Herzog [22] were the first to provide experimental evidence for type-II band alignment at the interface of a Si/SiGe heterostructure with tetragonal

strain distortion in both the layers. Abstreiter and his coworkers were the first to introduce the concept of a strain-induced type-II band alignment in Si/SiGe material system [23]. Experimental studies for the last few years on strained-SiGe materials have resulted in a significant progress in the understanding of strain relaxation kinetics and optimization of graded buffer layers with respect to relaxation and surface morphology [24, 25, 26, 27, 28]. These parameters are of crucial importance as they are interdependent and are affected by growth temperature, grading rate, and composition.

The sequence for fabrication of a strained-Si heterostructure involves three steps. First a relaxed graded-SiGe layer is formed to engineer the lattice constant of the SiGe alloy. The use of a relaxed-SiGe buffer layer to apply strain to epitaxial thin silicon films can be traced back to the work of Abstreiter et al. [23] in which a uniform $Si_{1-x}Ge_x$ buffer layer, thicker than the critical thickness, was grown at a high temperature. The strain in the buffer was partially relaxed and a thin Si layer was grown on top of it. The density of threading dislocations introduced as a result of the lattice relaxation was very high and resulted in low electron mobility at 4.2 K. The problem of high threading dislocation densities in relaxed structures may be avoided by using a series of low mismatched interfaces and increasing the Ge concentration in steps (step grading). The development of relaxed SiGe graded-layer technology provided an effective method for engineering the lattice constant of the SiGe film, and produced high electron mobility in strained-Si heterostructures.

The relaxed-SiGe graded layer is an epitaxial thin film with a sequence of layers that have a gradually increasing Ge content (typically the profile is linear) up to a final Ge composition. Use of compositionally graded layer to improve material quality of relaxed-SiGe buffer was first demonstrated by Fitzgerald et al. [29]. The compositional grading is believed to promote propagation while suppressing nucleation of dislocations. Xie et al. [30] have grown compositionally graded relaxed-$Si_{1-x}Ge_x$ buffer layers on Si with various composition gradients and temperatures. The authors reported a threading dislocation density in fully relaxed-SiGe buffer layers (see Fig. 3.6) grown using both molecular beam epitaxy (MBE) and rapid thermal chemical vapor deposition (RTCVD) in the range of $10^5$-$10^6$ cm$^{-2}$ [4].

Recent results have shown that the compositional grading can lower the threading dislocation density by three orders of magnitude and resulted in a much-improved electron mobility at low temperatures. In fact, the use of a compositionally graded, relaxed, $Si_{1-x}Ge_x$ buffer layer has been advocated as "virtual substrate" and allows the strain in the film to be tailored at will (in a biaxial/uniaxial fashion). Next a constant composition Ge film is grown to spatially separate the subsequent strained-Si film from the misfit dislocations that are intentionally introduced in the relaxed-SiGe graded layer. The final step is the deposition of the Si film, which is placed in a state of biaxial tension as it conforms to the lattice of the constant-composition SiGe layer. Once a low-defect density SiGe film can be fabricated, the strained-Si heterostructures can be tailored to enable various properties. One of the key determining

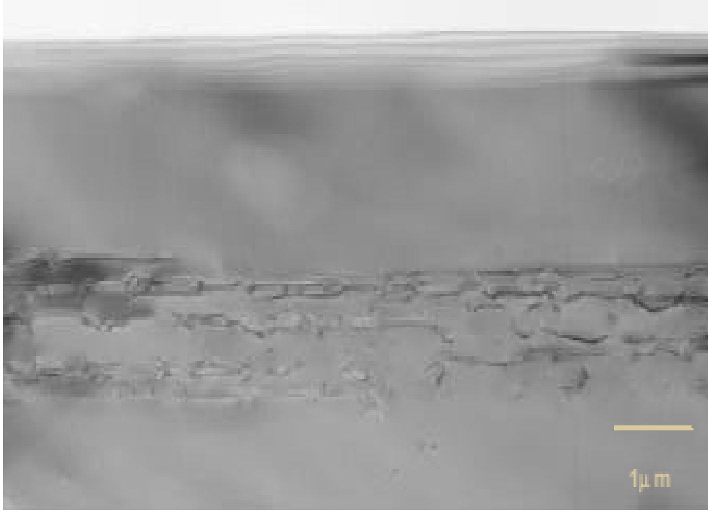



**FIGURE 3.6**
Confinement of threading dislocations. Source: IQE plc.

factors for the performance enhancement available with strained-Si technology is the amount of strain imparted to the strained-Si thin film, which is determined by the Ge content in the underlying SiGe film. Both electron and hole mobility enhancements increase with increasing strain. Using a higher Ge fraction in the underlying SiGe layer increases the degree of strain in the Si. However, hole mobility enhancement in strained-Si has a more complicated behavior, but can extend beyond the electron mobility enhancement achievable in strained-Si if enough strain is imparted to the film.

It has been shown theoretically and experimentally that the electron and hole mobilities increase with the Ge fraction in the relaxed-SiGe layer, but their dependence is different [31]. 15% Ge gives about 70% enhancement for electron mobility, and the enhancement approaches saturation beyond 15% Ge. Similar enhancement on hole mobility requires much higher Ge content and hole mobility starts to saturate at 35% of Ge fraction based on theoretical predictions. Although working with higher Ge content in the SiGe layer increases the performance enhancements of the nMOS and pMOS transistors, the tradeoff is that the strained-Si film needs to be within the critical thickness. It has been shown that in SiGe films, concentrations greater than 96% Ge, SiGe, and Ge nucleate higher levels of dislocations readily, increasing the dislocation density from $5\times10^5$ cm$^{-2}$ range to $\sim 10^6$ cm$^{-2}$. However, by grading Ge to 96% in the alloy and not continuing to 100%, threading dislocation density could be controlled in the $10^5$ cm$^{-2}$ range [32].

Many methods exist for depositing low-temperature silicon and $Si_{1-x}Ge_x$ on silicon. These can be broadly categorized into physical deposition and chemical vapor deposition (CVD) methods. On the chemical vapor deposition side, Gibbons and his group at Stanford were one of the first groups to demonstrate high-quality $Si_{1-x}Ge_x$ on silicon. Their lamp-heated limited reaction processing reactor (LRPCVD) laid the groundwork for other lamp-heated systems at Princeton, AT&T, and UT-Austin. The ultrahigh vacuum chemical vapor deposition (UHVCVD) reactor by Meyerson and his coworkers at IBM appeared at nearly the same time as LRPCVD. Combining a standard diffusion furnace with ultrahigh vacuum, they have made the most significant impact in the fabrication of Si/$Si_{1-x}Ge_x$ HBTs. Results of $Si_{1-x}Ge_x$ depositions at atmospheric pressure by ASM have been published. These atmospheric CVD results may be the most promising for widespread application of $Si_{1-x}Ge_x$ on silicon heterostructures in a production environment.

Gas source molecular beam epitaxy (GSMBE) [33, 34] have also been successfully employed for the growth of high-quality completely lattice-relaxed step-graded SiGe buffer layers on Si (001) in the temperature range of 750 and 800$^o$C. A more abrupt compositional transience of the SiGe/Si interface is expected in gas-source MBE-grown quantum wells, owing to reduced Ge segregation at the heterointerface [35], than in those grown by solid-source MBE where Ge segregation has been recognized as an important issue [36]. Another advantage of GSMBE is that uniform thickness and composition can be obtained without sample rotation. However, GSMBE is associated with autodoping of doping gas impurities, which would affect the device characteristics.

High-quality completely lattice-relaxed-SiGe buffer layers have been grown on Si (001) using molecular beam epitaxy in the temperature range of 750 and 900$^o$C and compositional grading on the order of 10% $\mu$m$^{-1}$ or less with final Ge concentrations of about 30%. Xie et al. [30] have grown compositionally graded relaxed-$Si_{1-x}Ge_x$ buffer layers on Si with various composition gradients. The authors reported a threading dislocation density in fully relaxed-SiGe buffer layers grown using both molecular beam epitaxy and rapid thermal chemical vapor deposition in the range of $10^5$-$10^6$ cm$^{-2}$ [4].

Relaxed $Si_{1-x}Ge_x$-on-insulator (SGOI) is a very promising technology, as it combines the benefits of two advanced technologies: the conventional SOI technology and the SiGe technology. The relaxed graded SiGe-buffer has allowed for the development of a multitude of new heterostructures with enhanced properties relative to bulk-Si. Early work focused on using these relaxed buffers as templates for inducing tensile strain in silicon channels. A possible extension of using SiGe buffers to engineer strained heterostructures involved the use of compressively strained Ge-rich SiGe layers that served as enhanced conduction pathways for holes, making them ideal for p-MOSFET applications [37]. By selectively utilizing the band alignments of strained-Si, strained-SiGe, and strained-Ge for optimal carrier confinement, the relaxed-SiGe buffer has been used to fabricate dual-channel heterostructures that

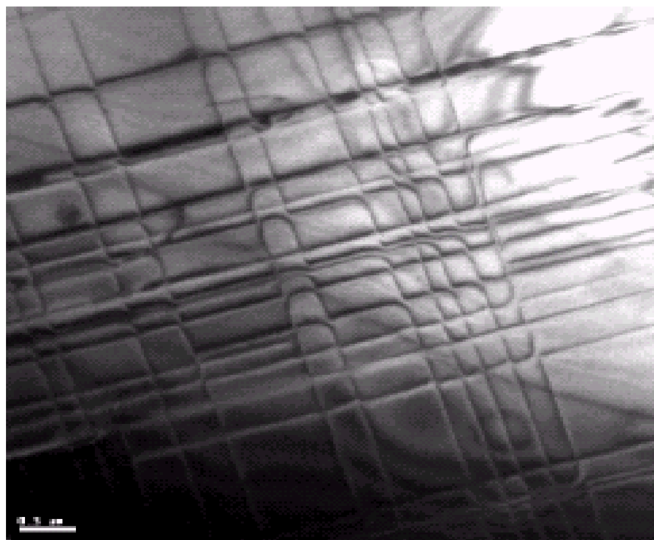


**FIGURE 3.7**
Planar view TEM image of a 21-nm thick strained-Si layer on $Si_{0.72}Ge_{0.28}$ after 30 s at 1000$^o$C. After P. M. Mooney, 2003 International Conference on Characterization and Metrology for ULSI Technology, March 24-27, 2003.

exploit the enhanced electron mobility of strained-Si as well as the enhanced hole mobility of strained Ge-rich layers [38, 39]. Dual-channel devices utilizing relaxed-SiGe buffers graded to $Si_{0.5}Ge_{0.5}$ have even exhibited nearly symmetric electron and hole mobilities [40].

The utility of the relaxed-SiGe buffer has been further extended to allow for the transfer of various scalable, lattice-mismatched layers to Si handle wafers. Successful demonstrations of layer transfer using relaxed-SiGe buffers include strained-Si on insulator (SSOI), SiGe-on-insulator (SGOI), germanium-on-insulator (GeOI) [41, 42, 43, 44]. As the crosshatch surface roughness inherent to efficient relaxation during the grading process in SiGe buffer (see Fig. 3.7) growth precludes the possibility of successful bonding, a chemical-mechanical planarization (CMP) step is required (see Fig. 3.8) to reduce to the RMS surface roughness (see Figs. 3.9 and 3.10) to a level suitable for bonding.

Above the critical layer thickness and Ge concentration, differences between the lattice constant of the SiGe layer and that of underlying bulk silicon cause misfit dislocations and related crystal defects. These defects cause surface roughness, which lowers the mobility and reduces circuit yield. To control dislocations, one can employ a compositionally graded series of layers of SiGe with varying Ge concentration underneath the SiGe buffer layer combined with CMP. By gradually expanding the crystal lattice beyond that of Si,

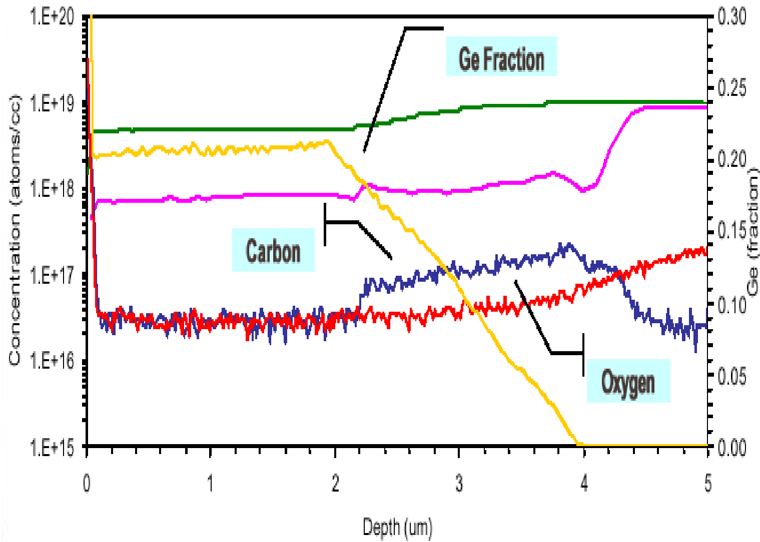



**FIGURE 3.8**
SIMS Profile of IQE-grown graded-SiGe layer showing high level of control. Source: IQE plc.

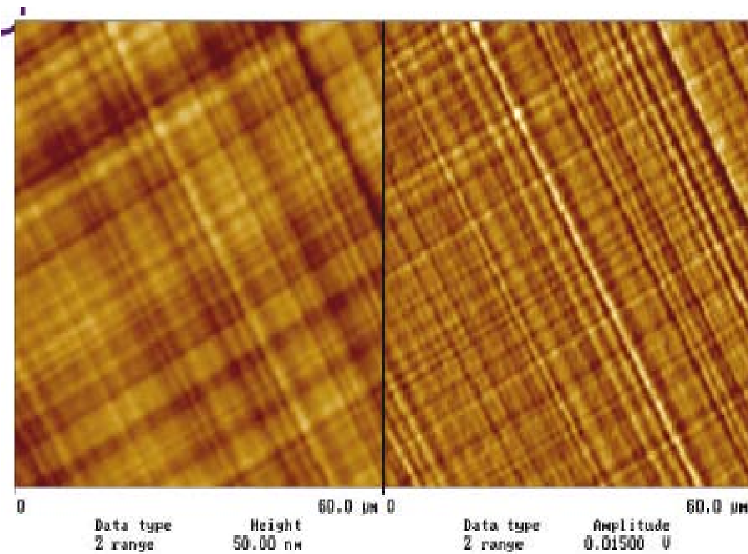



**FIGURE 3.9**
Surface microroughness prior to polish. AFM area: 60$\mu$m$\times$60$\mu$m, RMS value typically 6.5 nm and Ge = 20%. Source: IQE plc.

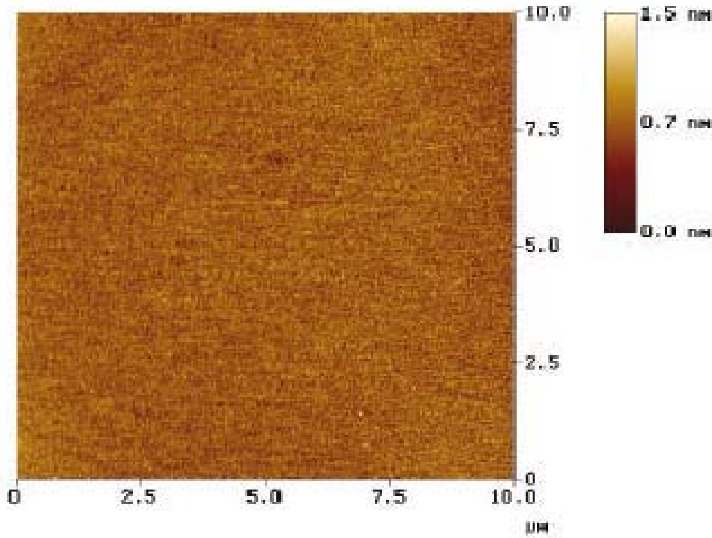



**FIGURE 3.10**
AFM image of polished buffer surface, RMS roughness = 0.13 nm, $R_a$ = 0.10 nm. Source: IQE plc.

dislocations are controlled and confined to the SiGe buffer layer.

It is important to recognize the importance of the relaxed buffer technology with wafer bonding. Wafer bonding and layer transfer are important techniques in heterointegration. However, it has been limited primarily by the following three factors:

• any desired lattice can be created and one is not limited to bulk substrate lattice constants: Thin strained layers can be created from any semiconductor material (for example, a SiGe alloy lattice constant is required to create thin layers of strained-Si),

• the bulk material of each bonded wafer is the same, and thus wafer separation or cracking due to thermal expansion differences during the bonding and annealing process are eliminated, and

• entire 200 and 300 mm wafers with any semiconductor on the surface can be created since the original lattice engineered material is on the same diameter substrate. This is very important, as it truly brings new functionality at low cost to a variety of applications.

### 3.3.2 Substrate Specifications

Strained-Si technology provides great performance benefits for CMOS, but as with any new technology that is introduced to a manufacturing infras-

**TABLE 3.2**
Typical specifications for strained-Si Substrates. Source: IQE plc.

| Parameter | Typical Value | Technique | Comments |
|---|---|---|---|
| Strained-Si Thickness | 175 Å ± 5% | Spectroscopic Ellipsometry | SE calibrated to cross-sectional TEM |
| Strained-Si Thickness Uniformity | $> 3.5\%$ | Spectroscopic Ellipsometry | SE calibrated to cross-sectional TEM |
| Ge in SiGe Layer | 0.17±0.01 | Spectroscopic Ellipsometry | SE calibrated to XRD |
| Relaxation of SiGe Layer | $> 95\%$ | X-ray Diffraction | Triple axis |
| Surface Micro-roughness | $R_a < 10$ Å | AFM | 10 $\mu$m×10 $\mu$m scan size Avg. 5 sites: Center, R/2 and R |
| Field TDD of SiGe Layer | $< 3E5/cm^2$ | EPD | EPD verified by planview TEM |
| Pile-up TDD of SiGe Layer | $< 3$ cm/cm$^2$ | EPD | EPD verified by planview TEM |
| LPD | $< 250@0.5\mu m$ | Tencor Surfscan | Resolution of defect analysis limited by surface roughness and pile-up TDD |
| Oxygen Concentration | $< 1.0E18/cm^{-3}$ | SIMS | Cs SIMS analysis |
| Carbon Concentration | $< 1.0E18/cm^{-3}$ | SIMS | Cs SIMS analysis |
| Metallic Surface Contamination | $< 3.0E10/cm^2$ | VPD | Fe, Cu, Cr, Zn and Ni $< 5.0E10/cm^2$ for Ca, Na and Ge |
| Doping | Customer Defined | SRP | Doping to customer requirements |

tructure, the ease with which it is adopted is very dependent on the number and severity of changes the new technology requires. For strained-Si technology, the key question for many manufacturers is whether the substrate technology can be taken from research laboratories, which typically use UHV epitaxy techniques, to large-scale production, which requires high throughput, low-cost techniques. Strained-Si processes have now been demonstrated and devices have been produced on strained-Si substrates fabricated using industry-standard low-pressure Si epitaxy systems, such as the ASM Epsilon and Applied Materials' Epi Centura platforms. Epiwafer manufacturer IQE Silicon Compounds has announced the commercial availability of strained-Si substrates (see Table 3.2).

### 3.3.3 Strained-Si on Insulator

Incorporation of SOI in advanced CMOS for high-performance applications has already taken place. The use of a buried oxide underneath the bulk-Si CMOS devices results in a number of benefits, including reduction of junction capacitances, increased circuit density, and reduced latch-up. If strained-Si is incorporated directly on an insulator, without the presence of SiGe in the final device structure, some of the problems encountered in SiGe, e.g., high leakage current, Ge diffusion, and enhanced n-type dopant diffusion can be avoided. Thus, there is a need to study the fabrication techniques and material properties of ultrathin strained-Si films on insulators.

The presence of a SiGe layer in virtual substrate causes material and process integration challenges. Using ultrathin strained-Si layer on SOI structures that bypass this SiGe layer, provides high electron mobility while eliminating material and process integration problems. Currently, research is being conducted in exploring the possibility of including biaxial strained-Si in mainstream CMOS process to enhance device performance. A novel approach in this direction is to fabricate ultrathin strained-Si layer on SOI, called strained silicon-on-insulator (SSOI), a combination of SOI technology with strained-Si, thereby complementing their individual advantages.

Various techniques have been employed for the formation of relaxed SiGe-on-insulator and strained Si-on-insulator [43, 45, 46, 47]. These techniques involve SIMOX of SiGe virtual substrate, Ge condensation, and wafer bonding of virtual substrates. Strained-Si layers on Ge-free structures have also been reported [41, 48]. In order to study strain behavior in ultrathin strained-Si on insulator films, Drake et al. [48] transferred a 130 Åthick strained-Si layer, to an oxidized handle wafer via wafer bonding and an etch-back process. Two SSOI structures were investigated for strain relaxation characterization. The first structure consists of a 130 Åthick strained-Si directly on 300 nm of $SiO_2$. The second structure consists of a 100 Åthick strained-Si with an underlying 8-nm thick SiGe layer on 300 nm of $SiO_2$. In Ge-free structures, the transfer SiGe layer has been omitted altogether, producing a strained-Si layer bonded directly to an oxide. The strain is still maintained by the bonds to the oxide

even after high temperature annealing, showing that the underlying SiGe is not needed for supporting the strain. Such a structure is important, due to the absence of Ge in the final heterostructure, allowing for conventional high-temperature Si processing that would not be possible in the presence of Ge in the bottom SiGe relaxed layer.

The process of fabricating SSOI structure involves transferring a strained-Si/relaxed-SiGe heterolayer on a handle wafer, followed by selective etch-back of SiGe to leave SSOI structure [49, 50]. A homogeneous stress of 1.5 GPa has been exhibited by Soitec [51] manufactured SSOI wafers with a 200 Å Si layer. The SSOI structure is fabricated by transferring strained-Si grown epitaxially, or layer by layer, on relaxed-SiGe to a buried oxide layer. The SiGe layer also may be removed before fabricating the device. Strain retention has been confirmed in the strained-Si layer after the layer transfer process and thermal cycles. Electron and hole mobility enhancements were confirmed in MOSFETs fabricated on SSOI. Fabrication of sub-60 nm FETs were also demonstrated on SSOI.

SSOI substrates without an intermediate SiGe or $SiO_2$ layer were fabricated by wafer bonding and layer transfer, as shown in the process flow diagram in Fig. 3.11. The process involved growth of relaxed-SiGe virtual substrate at $900^oC$ via ultrahigh vacuum chemical vapor deposition to fabricate relaxed-$Si_{0.76}Ge_{0.24}$ layer which was compositionally-graded at 10% Ge $\mu m^{-1}$ with a 3 $\mu$m $Si_{0.76}Ge_{0.24}$ cap layer. The combination of a low grading rate and a high growth temperature results in complete relaxation with threading dislocation densities of $\sim 10^5$ $cm^{-2}$.

Shown in Fig. 3.12 are cross-sectional transmission electron microscopy images of the SSOI structure (a) before and (b) after SiGe removal. The arrows shown in Fig. 3.12(a) indicate the location of misfit dislocations with line directions normal to the image plane. Note the average dislocation spacing of these dislocations is approximately 40 nm. The RMS surface roughness of the structure after mechanical grinding and KOH etching was determined via TM-AFM to be approximately 30 nm over a 25 $\mu$m $\times$25 $\mu$m area, with the cross-hatch pattern that is characteristic of relaxed-SiGe buffer layers reappearing due to the anisotropic nature of the KOH etch. Shown in Fig. 3.12(b) is the final strained-Si on Si structure, demonstrating the complete removal of the SiGe layer after grinding and etching. At the inset of Fig. 3.12(b) is a high-resolution TEM image of the bond interface, which indicates the lack of an intermediary oxide layer. The RMS surface roughness of the strained-Si after SiGe removal was approximately 1.9 nm for a 10$\mu$m $\times$10 $\mu$m TM-AFM scan, signifying excellent surface quality. These structures eliminate the SiGe layer before transistor fabrication, thereby providing higher mobility while mitigating the SiGe-induced material and process integration problems.

IBM reported strained-Si channel MOSFETs which are built directly on insulator structures with no SiGe layer present under the strained-Si channel. Without the relaxed-SiGe layer present as the lattice template, it is necessary that the insulator layer supports the strain state in the Si layer. SSOI offers

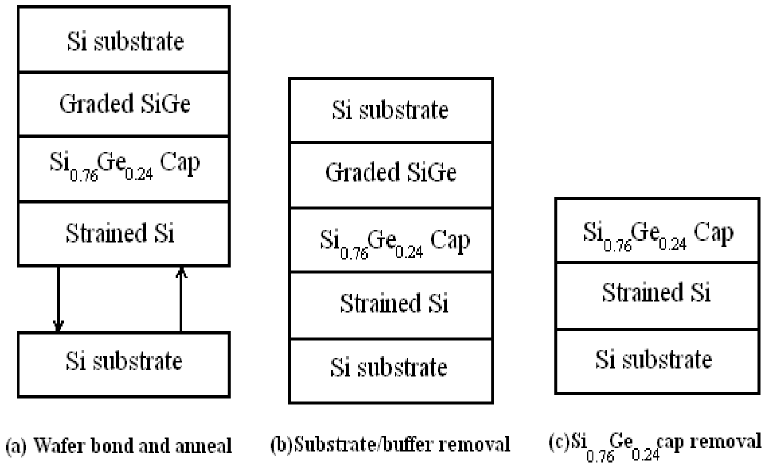



**FIGURE 3.11**
SSOI fabrication process showing (a) wafer bond and anneal, (b) structure after Si substrate and SiGe graded buffer removal, and (c) structure after removal of the $Si_{0.76}Ge_{0.24}$ cap. After D. M. Isaacson et al., SMA Symposium 2005.

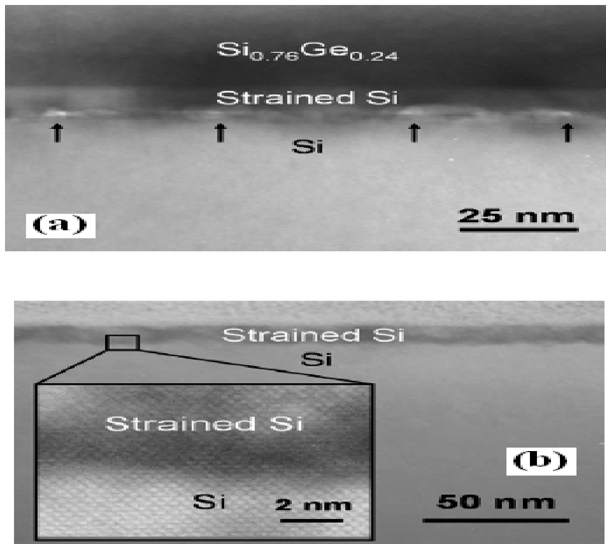



**FIGURE 3.12**
Cross-sectional TEM images (a) before and (b) after SiGe layer removal. Arrows in (a) indicate the location of misfit dislocations at the interface. After D. M. Isaacson et al., SMA Symposium 2005.

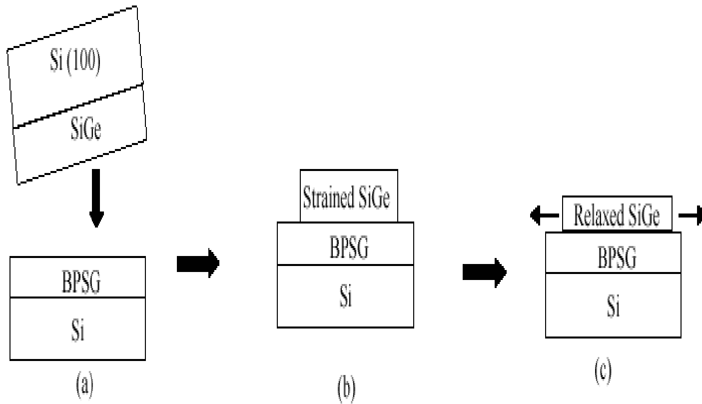



**FIGURE 3.13**
Process flow of relaxed-SiGe on BPSG insulator: (a) strained-SiGe grown pseudomorphically on Si(100) is transferred onto a BPSG layer by wafer bonding and Smart-cut processes, (b) strained-SiGe islands on BPSG are formed by plasma etching, and (c) strained-SiGe islands relax by lateral expansion as BPSG softens during high temperature annealing. After H. Yin et al., *IEEE IEDM Tech. Dig.*, 2003(53-56). With permission.

the possibility to scale to Si body and to form fully depleted devices on SOI. At the future advanced process nodes of 65 nm and below, performance and power consumption issues arise due to bulk silicon's higher leakage currents. SSOI technology has proven thus far to be a promising solution to deal with these issues.

To avoid the inevitable dislocations in conventional relaxed-SiGe buffers, Yin et al. [52] have developed a process for defect-free relaxed-SiGe on a BPSG insulator. Thin, fully strained SiGe layers without defects are first grown on a silicon substrate. The strained-SiGe is transferred by wafer bonding and Smart-Cut technique to a BPSG layer, and during annealing the SiGe can expand laterally and relax as the BPSG softens (Fig. 3.13).

Strained-Si could be grown on these layers, but the SiGe would be present during MOSFET fabrication and in the final structure. To obtain SiGe-free strained-Si on insulator, the process of Fig. 3.14 was developed, in which unstrained 25-nm Si was grown on top of 30-nm strained $Si_{0.7}Ge_{0.3}$ before transfer. The Si/SiGe was transferred onto BPSG films, resulting in a strained-SiGe/unstrained-Si/BPSG structure. The SiGe/Si film was then patterned into islands of edge width from 30 to 200 $\mu$m. Annealing at 800$^o$C softened the BPSG and allows the SiGe and Si stack to coherently expand, resulting in same strain change in both films. The compressive strain in the SiGe relaxes and generates tensile strain in the Si which was confirmed by Raman

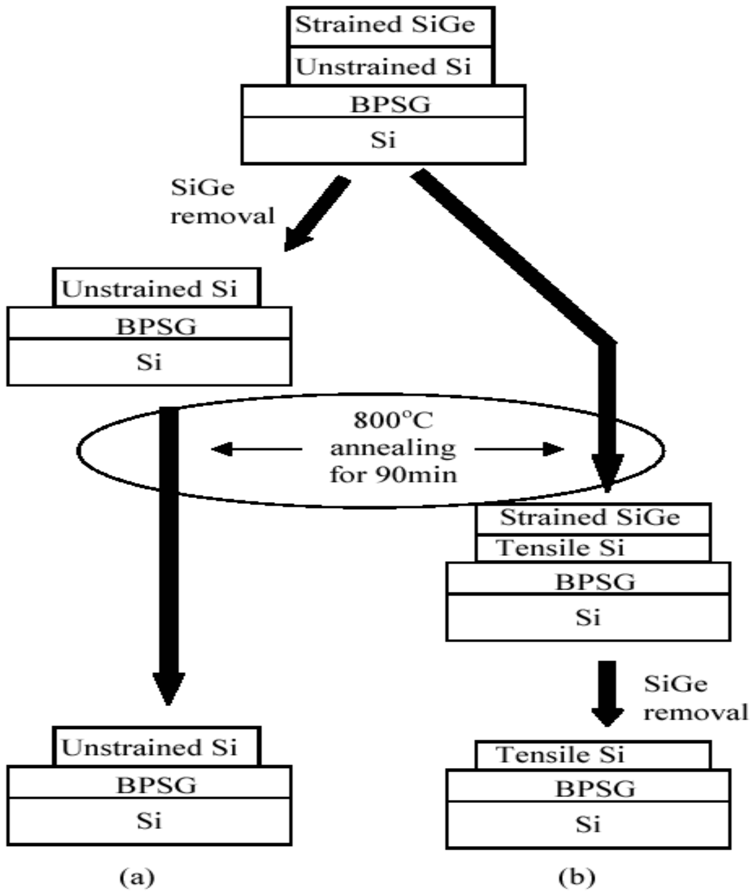



**FIGURE 3.14**
Process flow of (a) unstrained-Si on a BPSG insulator and (b) strained-Si on a BPSG insulator. The strained-SiGe/unstrained-Si/BPSG at the top of this process flow diagram is made by the same process as shown in Figs. 3.13(a) and (b), except that a thin unstrained-Si layer is grown on top of the SiGe before transfer. After H. Yin et al., *IEEE IEDM Tech. Dig.*, 2003(53-56). With permission.

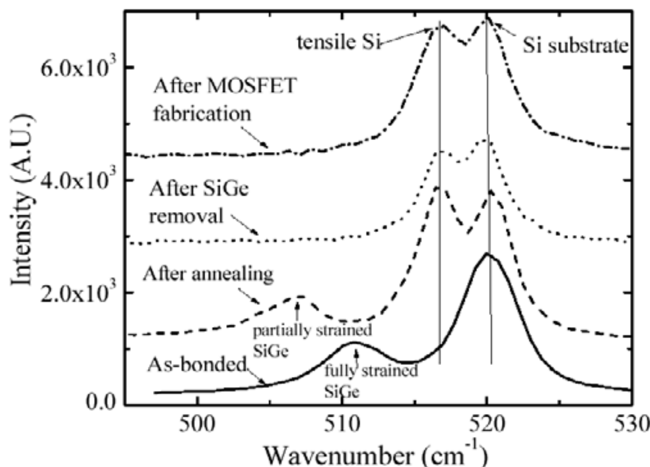



**FIGURE 3.15**
Raman spectrum (Si-Si nearest neighbor vibration mode) measured at the center of a 90 $\mu$m$\times$90 $\mu$m island at various stages of processing. The peak positions depend on composition and strain. The annealing reduces compressive strain in the SiGe and adds tension in the Si. The FET process does not alter the strain in the Si. After H. Yin et al., *IEEE IEDM Tech. Dig.*, 2003(53-56). With permission.

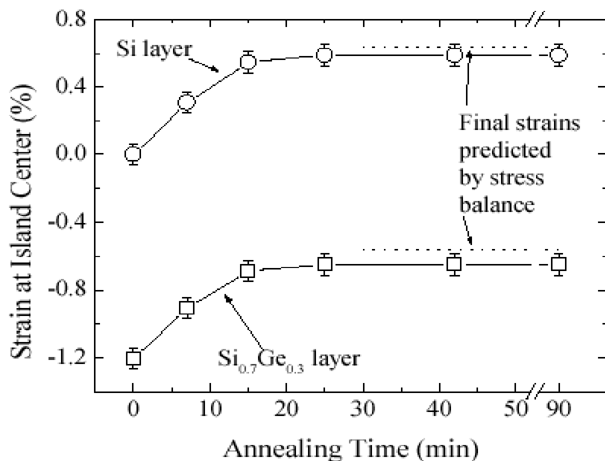



**FIGURE 3.16**
Biaxial strain (from micro-Raman spectroscopy) of 30-nm $Si_{0.7}Ge_{0.3}$ and 25-nm Si films at the center of a 30 $\mu$m$\times$30 $\mu$m island as a function of annealing time at 800$^o$C, showing evolution of strain in silicon. Dashed lines are calculations of stress balance. After H. Yin et al., *IEEE IEDM Tech. Dig.*, 2003(53-56). With permission.

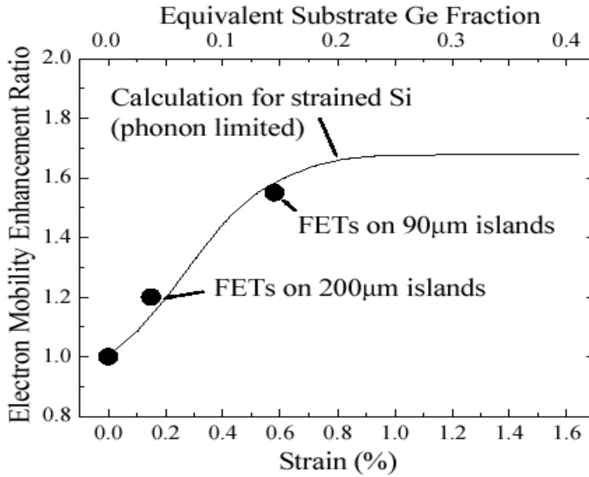



**FIGURE 3.17**
Electron mobility enhancement as a function of strain in the silicon channel layer. Also shown on the top x-axis is the Ge fraction in relaxed-SiGe buffers in conventional strained Si on relaxed-SiGe structures to give the same strain as that on the lower x-axis. After H. Yin et al., *IEEE IEDM Tech. Dig.*, 2003(53-56). With permission.

spectroscopy (see Figs. 3.15 and 3.16). The final strain in the SiGe and Si films upon equilibrium is governed by stress balance between the layers. The effective electron mobility near turn-on is found to be 55-60% higher in the strained-Si devices than in the control, in good agreement with the strain-induced mobility enhancement predicted by theory (Fig. 3.17) [53]. The presence of a SiGe layer causes material and process integration challenges. Using ultrathin SSDOI structures that bypass this SiGe layer, provides high electron mobility while eliminating material and process integration problems.

### 3.3.4 Smart-Cut

High-mobility strained-Si is now being incorporated into conventional CMOS, and future generations may incorporate both the benefits of strained-Si and silicon-on-insulator. One substrate that can enable the above consists of low-defect, monocrystalline, relaxed SiGe-on-insulator with a surface layer of strained-Si, yielding a strained-Si-on-insulator structure. Equivalently, SSOI substrates can also be fabricated such that the strained-Si lies directly on the insulating substrate, with no underlying relaxed-SiGe layer.

Separation by implanted oxygen (SIMOX) is a popular technique for SOI fabrication, but unfortunately, is only applicable to SGOI fabrication for low Ge mole fractions. The technique involves the implantation of oxygen ions to a desired depth, followed by a high-temperature anneal to remove implant

damage from the crystalline layer, and to form the buried oxide layer. Fukatsu et al. [42] have successfully applied the approach to SGOI fabrication for a Ge content of 18%. Another method demonstrated by Tezuka et al. [54, 55] requires high-temperature oxidation of a low-Ge content SiGe layer on a thinned SOI substrate. The oxidation leads to the expulsion of Ge ahead of the oxide front, and the high temperature causes Ge interdiffusion into the underlying thinned Si. As the strain relaxation mechanism for the Ge-rich layer is lattice expansion and not dislocation nucleation, the buried oxide flows at a temperature. This approach produces moderately high defect densities when applied to blank wafers, but yields very low defect densities for patterned mesas [55].

In addition to the above approaches, wafer bonding is another technique for fabricating SGOI or SSOI, involving the bonding of wafers with SiGe films to insulating substrates. A few widely employed techniques exist for accomplishing layer transfer of a thin film of material onto any desired substrate: (a) back-side grinding, (b) grind and etchback, and (c) delamination via implantation. In the grind and etchback approach, wafers are bonded, the back side of the wafer containing the transfer layer is thinned by grinding, possibly followed by an etch to remove the remaining excess material leaving the transfer layer. The other method, delamination via implantation, employs ion implantation prior to wafer bonding, followed by annealing to cause delamination and layer transfer at the implant depth. In these methods, a final polishing step is usually employed to smooth and thin the transferred layer. Huang et al. [44] used Smart-Cut to transfer relaxed-SiGe onto oxidized Si substrates, followed by CMP to thin and smooth the transferred layer. The limitation of the above methods is the lack of control over the final SiGe thickness on the handle wafer, and the wafer-scale uniformity of the transferred layer. Two methods for fabricating monocrystalline, relaxed-SiGe and strained-Si layers on insulating handle substrates have been reported [56] which involve (a) back-side grinding and etchback, and (b) delamination via implantation. These processes were combined with a stop layer, allowing for the fabrication of transferred layers with uniform thickness across the wafer.

The Smart-Cut approach has several advantages: (a) no etch-back step, (b) the prime-quality wafer; and it is fully recyclable, and UNIBOND reduces to a single-wafer process, (c) only conventional equipment is needed for mass production, (d) relatively inexpensive 12 in. wafers are manufacturable, and (e) unlimited combinations of BOX and film thicknesses can be achieved in order to match most device configurations (ultrathin CMOS or thick-film power transistors and sensors).

The defect density in the film is very low, the electrical properties are excellent, and the BOX quality is comparable with that of the original thermal oxide. It is worth noting that the two interfaces of the BOX are ideally organized: the top interface (film-BOX) has the high quality expected from thermal oxidation whereas the bonded interface, of poorer quality, is located underneath the BOX and has little influence on the SOI device performance.

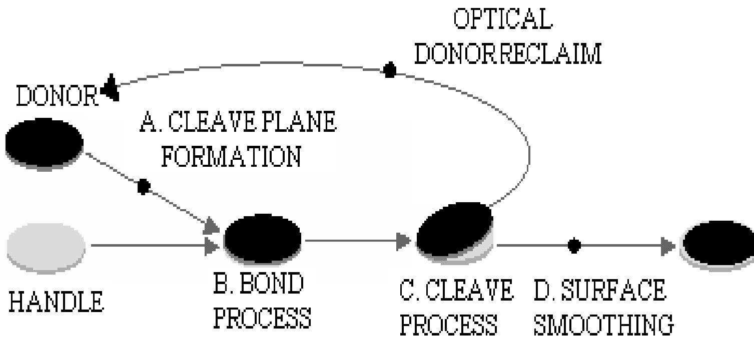



**FIGURE 3.18**
Generic layer transfer process flow. Source: SiGen.

A fascinating aspect is that the Smart-Cut process is adaptable to a variety of materials: SiC or III-V compounds on insulator, silicon on diamond or glass, etc. The possibility to enroll, in the SOI-based microelectronics, these materials with large band gap, photonic, or high-temperature capabilities opens exciting prospects for the integration of totally new types of devices.

### 3.3.5 Hybrid Substrates

Layer transfer (LT) is another technique for producing engineered substrates. In general, the engineered substrate can be designed by implementing layer transfer, in which the top layers are transferred to another substrate. Recent advances in layer transfer technologies allow ultrathin SOI and silicon-on-silicon substrates to be manufactured with the minimum allowable defects and device silicon-layer uniformity specifications for high-yield, deep-submicron manufacturing. The main steps include cleave plane formation, plasma bonding, mechanical cleave, and noncontact surface epi-smoothing. This LT process is flexible enough to allow SOI, multilayer stacked layers, strained film formation, as well as hybrid substrates such as (110) on (100), for example.

The use of layer-transfer methods for the fabrication of engineered substrates for semiconductor and optoelectronic applications have been proposed for decades. Fig. 3.18 depicts a generic layer-transfer method where the key technology differentiators are (a) how the cleave plane is formed, (b) how the wafers are bonded, (c) how the bonded wafers are cleaved, and (d) how the exposed film is smoothed or polished. These bond/cleave techniques include epilayer separation, controlled cleaving along a stressed layer, water-jet porous silicon layer separation, and hydrogen-induced thermal separation.

Hybrid-orientation technology (HOT) involves building standard CMOS on a hybrid silicon substrate with different crystal orientation in each of two lay-

ers. The layer transfer process, block-level trench etch, and epitaxy regrowth are performed before the conventional CMOS device process. For pMOS, hole mobility is known to be 2.5 times higher on (110) surface-orientation compared to that on standard wafer with (100) surface-orientation. IBM has created a hybrid-orientation technology where CMOS is fabricated on hybrid substrate with different crystal orientations to achieve significant pMOS performance enhancement. In the HOT technology, layer transfer process, block-level trench etch, and epitaxial regrowth were performed before conventional CMOS device process. An enhancement of 40-65% for the pMOS was demonstrated on a 90-nm node CMOS technology.

IBM has announced a technique to utilize different crystal orientations for pMOS (110 oriented surface) and for nMOS (100 oriented surface) to increase the mobility of the holes and electrons, respectively. Epitaxial regrowth from the substrate underneath the buried oxide is used to create the second orientation at the surface. CMOS devices with substantial enhancement of pMOS drive currents have been demonstrated at $L_{poly}$ = 80 nm and with physical gate oxide thickness of 1.2 nm.

### 3.3.6 GeOI Substrates

In general, in semiconductors, it is observed that if the lattice constant increases and elastic modulus decreases, the band gap decreases although the electron mobility increases. Thus, it is desirable to combine larger lattice constant materials on Si from the carrier mobility enhancement perspective. However, the lattice-mismatch between the desired film and silicon is large enough that many defects are created in thin films at the growth temperature, and therefore, the material quality is generally inferior for most applications. Also, low band gap, poor quality oxides, and other surface properties of engineered substrates have led to very limited application.

For sub-32 nm technology nodes there exists strong interest in exploring the use of Ge substrates instead of Si, mainly because of the enhanced carrier mobility and the overall lower thermal budget required for processing. This implies that some of the fundamental research performed on Si will have to be repeated for Ge. There is uncertainty whether or not a real push forward will be achieved with Ge substrates or with Germanium-on-Insulator (GeOI). The important aspects of Ge substrate technology are generally: (a) Bulk Ge substrates, (b) Ge layers transferred to oxidized Si handle wafers, thereby forming Ge-on-Insulator (GeOI), and (c) Epitaxial deposition of Ge thin films on Si substrates in various configurations. Bulk-Ge substrates (grown via the Czochralski method) are more difficult to make than bulk-Si substrates due to the lower critical resolved shear stress and lower mechanical strength of Ge. Although 300-mm dislocation-free Ge substrates can be fabricated, the economics of high-volume manufacture are still unsure. Development is still required to obtain site flatness and crystal-originated pit specifications that are equivalent to state-of-the-art Si substrates.

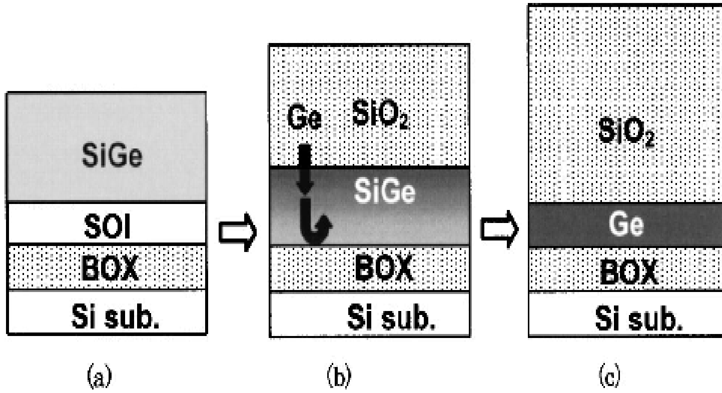



**FIGURE 3.19**
GeOI fabrication process by the Ge condensation by oxidation technique. (a) SiGe layer is grown epitaxially on an SOI wafer, (b) Oxidation of SGOI, and (c) Ge is condensed completely. After T. Tezuka et al., *Appl. Phys. Lett.*, Vol. 79, 2001(1798-1800). With permission.

The GeOI layer was fabricated by oxidizing a SiGe layer which was grown epitaxially on an SOI substrate. This fabrication method is called the Ge condensation technique [46]. This method has been also been applied to fabrication of a strained SiGe-on-insulator layer [57]. By further condensation of Ge atoms in a SGOI layer by oxidation at temperatures lower than the melting point of $Si_{1-x}Ge_x$, strained GeOI layers thinner than 10 nm have been fabricated. The fabrication process of GeOI layers is illustrated schematically in Fig. 3.19. At first, a $Si_{1-x}Ge_x$ layer with the Ge fraction of 15% is grown epitaxially by ultrahigh vacuum chemical vapor deposition on a (100) commercially available bonded SOI wafer, as shown in Fig. 3.19(a). Next, thermal oxidation in $O_2$ atmosphere was carried out (Fig. 3.19(b)) at temperatures lower than the melting point of $Si_{1-x}Ge_x$, which depends on the Ge fraction. In the oxidation process, Ge atoms were rejected from the $SiO_2$ layers, while outdiffusion of Ge atoms is suppressed by the buried and top oxide layers. Therefore, the total amount of Ge in the SiGe layer is preserved through the oxidation process. The SiGe and Si layers are merged into a uniform SiGe layer by interdiffusion between Si and Ge atoms. As a result, the Ge fraction in the SiGe layer increases when the SGOI thickness becomes thinner than the initial SiGe thickness. Finally, oxidation is stopped when almost all Si atoms in the SGOI layer are oxidized. Consequently, an ultrathin GeOI layer is obtained, as illustrated in Fig. 3.19(c).

Managing the lattice mismatch and other material property differences is the key to creating new engineered substrates. Fabrication of engineered

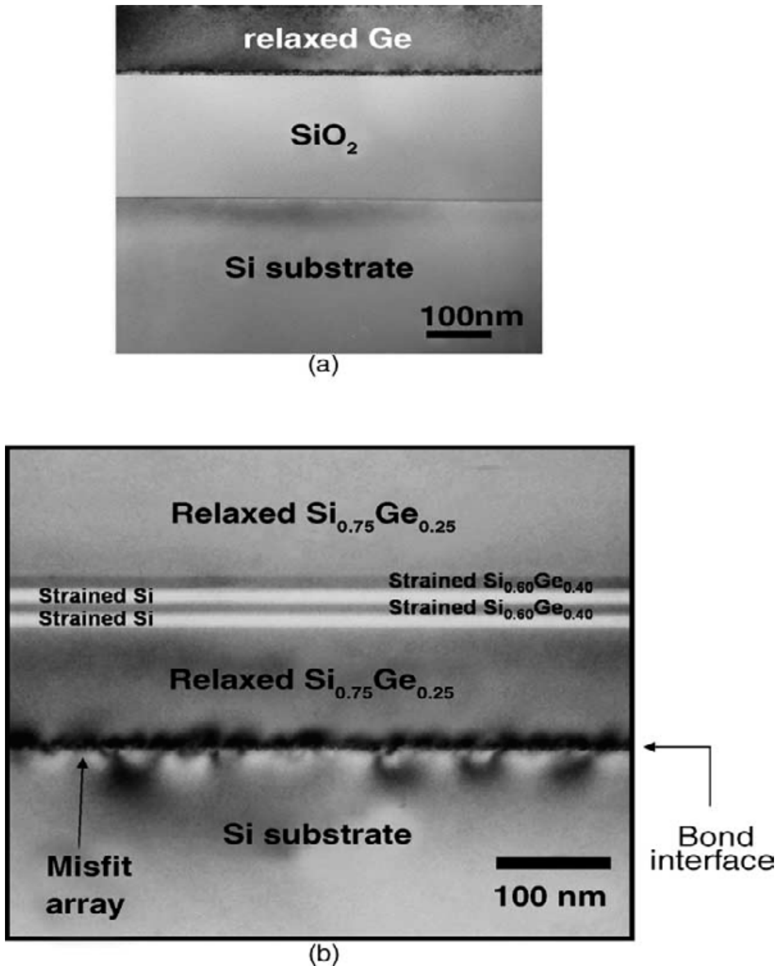



**FIGURE 3.20**
TEM cross-sections of structures which embody the flexibility of current substrate engineering methods: (a) Ge-on-insulator created from the transfer of a relaxed-Ge layer with low threading dislocation density to a $SiO_2$/Si wafer; (b) strained-Si/strained-SiGe multilayer on SiGe on Si (SGOS), in which the graded SiGe buffer layer and original Si substrate have been removed, improving thermal conductivity of SiGe-based MOSFETs. After A. J. Pitera et al., *J. Electrochem. Soc.*, Vol. 151, 2004(G443-G447). With permission.

substrates that can be used in silicon manufacturing, has to meet many challenges. Figs. 3.20(a) and (b) are some examples of engineered substrates that are now being fabricated via lattice engineering, planar strained layers, and layer transfer techniques. Fig. 3.20(a) is an example of Ge on insulator, potentially useful in integrated optoelectronic applications and also some high-mobility electronics platforms [58]. Fig. 3.20(b) shows the extent to which engineered substrate work is beginning to reveal the wide variety of structures that are now possible. Fig. 3.20(b) is a cross-section TEM micrograph of a SiGe/strained multilayer structure in which the original SiGe graded layer has been removed.

As it is unlikely that the complete replacement of bulk-Si substrates via Ge, it is expected that Ge substrate technology developments will be aimed at making Ge substrates that are suitable as donor wafers for repeated Ge exfoliation processes in the formation of Ge films on insulator (GeOI). Epitaxial Ge technology is especially promising for the implementation of strained Ge channels. GeOI technology addresses both the substrate manufacturability concerns and also reduces the likelihood of excessive Ge junction leakage during device implementation.

## 3.4 Characterization of Strained Layers

As discussed above, biaxial stress in a thin silicon layer (below the critical thickness for a given stress) is typically achieved by growing Si on a thick $Si_{1-x}Ge_x$ alloy (20-30% Ge, 1-5 $\mu$m thickness) as a virtual substrate. Alternatively, to combine the benefits of strained-Si with SOI technology, strained-Si can be grown on relaxed-$Si_{1-x}Ge_x$ on insulator using compliant substrates, SIMOX via condensation, or wafer bonding.

Although epitaxy and SiGe graded-layer processes are at the core of strained-Si technology, the manufacturability of devices is largely dependent on the structural properties of the substrates. The SiGe graded-layer epitaxial sequence typically results in a rough, cross-hatched surface morphology with a lateral periodicity of the order of 1 $\mu$m, and an RMS roughness greater than 4 nm. This rough wafer surface severely impacts the manufacturing of integrated circuits with state of the art ($< 0.5$ $\mu$m) lithography requirements, in terms of yield, in-line metrology, and photolithographic reliability. Thus a chemical-mechanical polishing (CMP) process is introduced between the epitaxy of the relaxed-SiGe buffer layer and the strained-Si. The CMP step eliminates the cross-hatched SiGe surface roughness, producing strained-Si wafers that largely appear like bulk-Si.

In all cases, critical parameters of the Si layer (stress, thickness, roughness, defect density, especially threading dislocations) need to be measured

carefully to achieve a manufacturable technology. If a SiGe buffer layer is used, one also needs techniques to determine Ge content (including its depth profile), buffer thickness, crystallinity (mosaic structure), interface roughness, and density and type dislocations and other defects. Nondestructive characterization of strained-Si structures is necessary to understand the thermal stability of strained-Si on SiGe.

Structural characterization is carried out by selective chemical etching, transmission electron microscopy (TEM), Rutherford Back Scattering (RBS), and double-crystal x-ray diffraction (DCXRD). Sputter depth profiling by secondary ion mass spectroscopy (SIMS) was used for chemical analysis. Electrical characterization included spreading resistance analysis on beveled cross-sections and four-point probe measurements.

We discuss a plethora of experimental techniques that have been used to characterize such strained-Si layers and their buffers and discuss their capabilities and issues. Techniques generally found especially useful are visible or UV Raman spectroscopy (to determine the stress in the Si channel), x-ray reflectivity (Si channel thickness), etch pit density (threading dislocation density), and transmission electron microscopy (thickness, misfit, and threading dislocation density, interface roughness, and other defects). Use of triple-axis high-resolution x-ray diffraction with synchrotron radiation has also been reported in the literature. Spectroscopic ellipsometry is normally the technique of choice for fast in-line thickness and composition measurements. However, significant errors may occur in the determination of thickness and Ge content in the buffer, if the change in the Si optical constants due to stress is neglected. It has been suggested in the literature that much more success can be achieved in determining the Si thickness with x-ray reflectivity, which is not affected by stress.

For Ge profile determination, a variety of surface science techniques (secondary ion mass spectrometry and Auger electron spectrometry) are used. The two critical factors here are to avoid charging by the ion beam and depth resolution. For very thin layers, one needs low primary ion energies, while high sputtering rates (and thus higher ion energies) are needed for thick $Si_{1-x}Ge_x$ buffers. One needs to avoid charging for thin electron-hole pairs in the $Si_{1-x}Ge_x$ layer and thus enhancing the conductivity. Meanwhile, it has been found that a high-energy electron beam coupled with Cs ion beam sputtering provides the best conditions to avoid charging issues with thick $Si_{1-x}Ge_x$.

Strained-Si substrates require careful characterization. X-ray diffraction and reflectivity are robust techniques and are sensitive to: parallel strain, perpendicular strain, tilt, twist and thickness. In the x-ray reflectivity (XRR), the fringe spacing is used to determine thickness. From this measurement, one can also determine density and roughness and this technique is more sensitive to thin layers than XRD. Reciprocal space mapping precision is typically 12 arc sec with a strain measurement accuracy of $\sim 1\%$. Measurements can be made within a short time. Thickness measurement accuracy of reflectivity

**TABLE 3.3**
Metrology needs for strained-Si layer characterization.

| | |
|---|---|
| SIMS | Ge profile, oxygen and carbon contamination |
| XRD | Ge mole fraction, triple axis for relaxation |
| AFM | Surface microroughness for polished layer and final device layer |
| TEM | Strained-Si thickness/uniformity, dislocation confinement |
| EPD | Schimmel Etch for TDD and pile-up density measurements |
| VPD | Surface metallic contamination |
| SE | In-line measurement of strained-Si thickness |
| SPV | Fe contamination levels |
| SRP | Dopant profile |

techniques is typically within 1%. For strained layers as the rocking curves are not very accurate, measurements with four (224) reciprocal space maps are used to eliminate effect of tilt. While the horizontal offset gives parallel lattice constant, the vertical offset gives perpendicular lattice constant. Combination of the lattice constant information gives parallel and perpendicular strain and composition. Tilt and twist are measured with x-ray diffraction. Bonded substrates have appreciable tilt and twist, since layer and substrate are formed from different crystal lattices, while bulk strained-Si and SIMOX SOI have negligible tilt and twist since layer and substrate are formed from the same crystal lattice. Raman spectroscopy is sensitive to parallel strain only. Additionally, FTIR and oblique incidence laser-scatter particle detection are also useful. Metrology needs for the characterization of strained layers are shown in Table 3.3.

Structural characterization and film quality of strained layers are usually carried out by transmission electron microscopy (TEM, both plain-view and cross-section) for the determination of defects/dislocations within the layers and to measure the thickness of the layers, and energy dispersive spectroscopy (EDS) utilizing a scanning TEM for film composition. Heterolayer surface roughness is another important parameter which adversely affects the performance and reliability of the devices. Surface morphology of buffer layers are dependent on the processing conditions such as Ge grading rate and growth temperature (see Fig. 3.21). While it has been shown that high growth temperature is desirable for dislocation propagation, it also leads to a 3D growth and eventually a rough surface morphology that may affect lithography steps in subsequent device processing. Atomic-force microscopy yields information about the surface morphology of a layer. In addition to AFM, Nomarski optical interference micrographs can also yield similar information about the surface morphology of a relaxed layer [59].

Hsu et al. [27] have investigated the surface morphology of a relaxed, com-

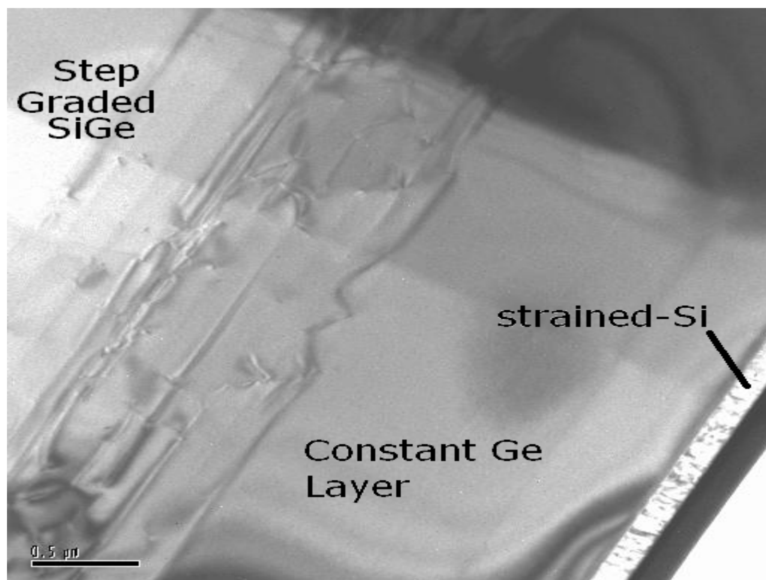



**FIGURE 3.21**
TEM of a UHVCVD-grown strained-Si layer grown at 800$^o$C on a step-graded relaxed-SiGe buffer. After C. K. Maiti et al., *Solid-State Electron.*, Vol. 48, 2004(1369-1389). With permission.

positionally graded $Si_{1-x}Ge_x$ structure, to illustrate the influence of defect-related strain fields on film growth. Quantitative topographic measurements via scanning force microscopy show the roughness associated with the cross-hatch patterns oriented along the (110) direction, which basically results from a local accumulation of the underlying misfit dislocations beneath the surface.

Atomic-force microscopy was used for acquisition of high resolution 3D topographs over an extended scale. The digitized surface topography is represented by a 2D array whose elements express the height information of each point of the 2D scanned image grid. As shown in Fig. 3.22, ULPCVD-grown strained-Si films show a root mean square (RMS) surface roughness and an average roughness ($R_a$) of 3.1 Å and 2.45 Å, respectively.

### 3.4.1 XRD

X-ray diffraction (XRD) allows the direct calculation of the Ge content and the degree of relaxation for heteroepitaxial SiGe films. For pseudomorphic or fully relaxed films the Ge content can be calculated from a simple $\theta$-$2\theta$ scan of the symmetric 004 reflection. The 004 reflection measures the crystal lattice spacing in the direction of the surface normal, i.e., the out-of-plane lattice spacing. Fig. 3.23 is a typical rocking curve obtained with a Philips X'Pert PRO MRD system to study the lattice strain in SiGe layers.

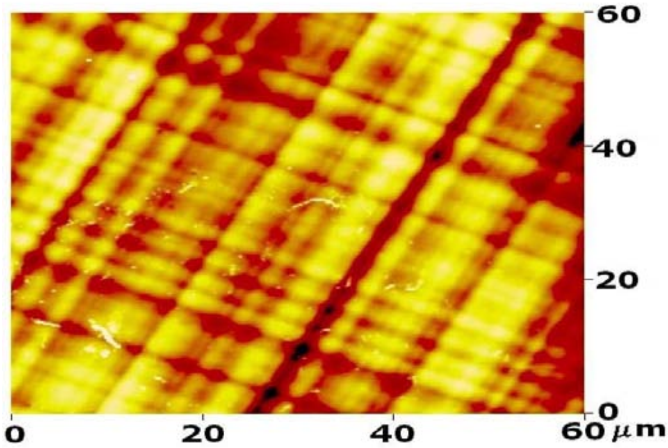



**FIGURE 3.22**
AFM surface of a ULPCVD-grown strained-Si layer grown at 800$^o$C on a step-graded relaxed-SiGe buffer. After C. K. Maiti et al., *Solid-State Electron.*, Vol. 48, 2004(1369-1389). With permission.

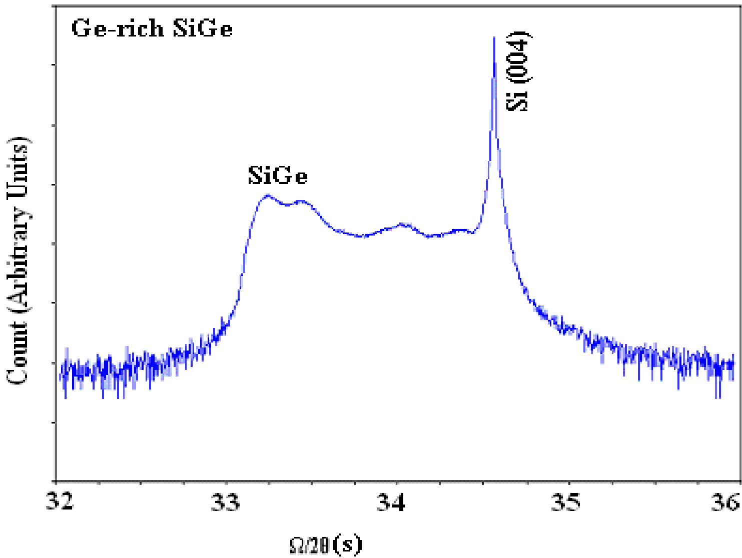



**FIGURE 3.23**
High-resolution x-ray rocking curve around Si (004) reflection for as-grown SiGe heterolayers.

For partially relaxed films additional information from an asymmetric peak (usually 224) is necessary in order to separate the influence of relaxation and Ge content on the lattice constants and thereby the peak position. The 224 reflection measures the spacing of lattice planes at an angle to the surface normal, i.e., a set of planes whose lattice spacing has both an out-of-plane and an in-plane component. The composition and the strain are calculated from the angular separation between the SiGe layer and the Si substrate peaks in the two measurements. For pseudomorphic growth, the 224 reflections of the different layers in a reciprocal space map (RSM) are located vertically above each other, so that the SiGe peak is found directly underneath the Si substrate reflection. With increasing relaxation the layer peak is shifted into direction of the line connecting the origin of the RSM with the substrate peak.

For full relaxation, the peak is located directly on this line. However, a deviation may occur in case of a possible tilt between the SiGe layer and the Si substrate, which can be accounted for by additional recording an RSM of the symmetric 004 reflection. For the calculation of Ge content and degree of relaxation of the constant concentration SiGe buffer the elastic constants of SiGe need to be interpolated linearly between Si and Ge, whereas a parabolic approximation is used for the dependency of the Ge content on the relaxed-SiGe lattice constants. The thin strained-Si layer can also be observed in the RSMs. In the 224 RSM, the respective diffraction pattern is located directly above the SiGe peak, i.e., the Si is pseudomorphic (fully strained) with respect to the SiGe buffer with a common in-plane lattice constant.

Fig. 3.24(a) shows the 004 data for a 21-nm thick Si cap layer on a $Si_{0.72}Ge_{0.28}$ buffer layer taken at two different regions, one with and one without the strained-Si cap layer. As is clearly seen when comparing the two scans, the features to the right of the Si substrate peak are from the strained-Si cap layer. Although the Si layer thickness can be determined from the solid curve in Fig. 3.24(a), comparison with the simulation is more precise when the two scans are subtracted. The difference data from the two scans in Fig. 3.24(a) is shown in Fig. 3.24(b), along with a simulation. The thickness of the Si cap layer determines the spacing of the oscillations and the strain determines their position with respect to the Si substrate peak. The shift of the experimental data from the Si cap layer with respect to the Si substrate peak is clearly visible. With this method of analysis the strain in the Si cap layer can be determined to within 1 or 2%, depending on the sample.

Fig. 3.25 shows the measured RSMs of both the symmetrical (004) and asymmetrical ($\bar{2}\bar{2}4$) diffraction regions of reciprocal lattice points (RLPs) for the $Si_{1-x}Ge_x$/Si heterolayers. The scans were performed along $\omega/2\theta$ direction (i.e., radial from the origin (000) of reciprocal space) for different $\omega$-offsets (i.e., transverse along a circle with center (000)) using the proper transformation from angular space into reciprocal space. In order to get intensity contour plots and to determine the Ge content and the relaxation status of the SiGe layers from reciprocal space maps, diffractions along (004) and $\bar{2}\bar{2}4$ directions

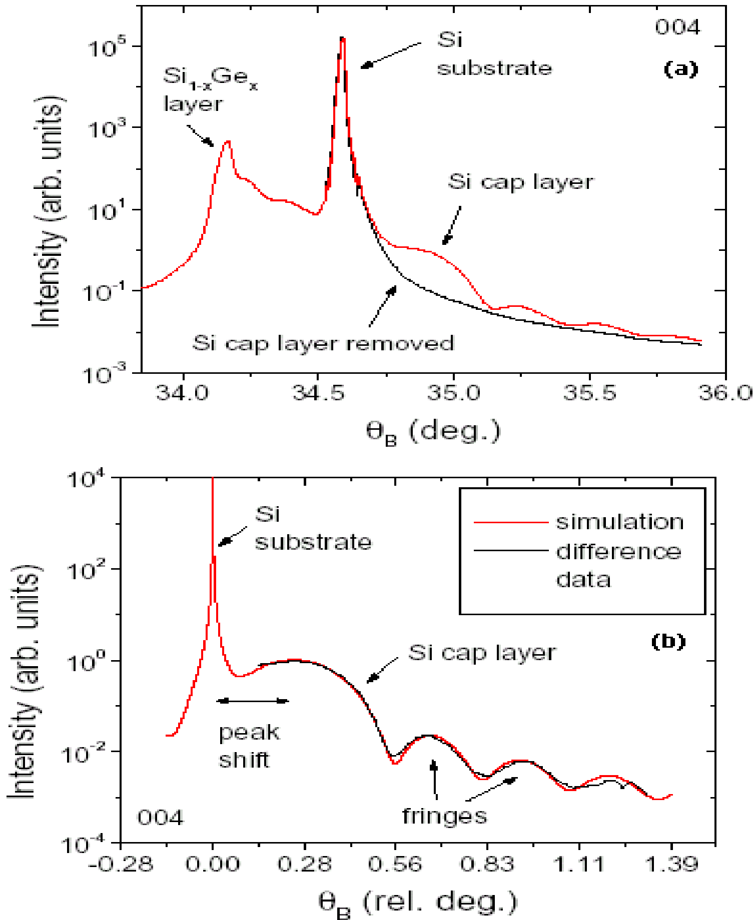



**FIGURE 3.24**
(a) 004 x-ray scans taken at regions with and without the 21-nm thick strained-Si cap layer on $Si_{0.72}Ge_{0.28}$ and (b) the difference between the two scans plotted with the simulation. After P. M. Mooney, 2003 International Conference on Characterization and Metrology for ULSI Technology, March 24-27, 2003.

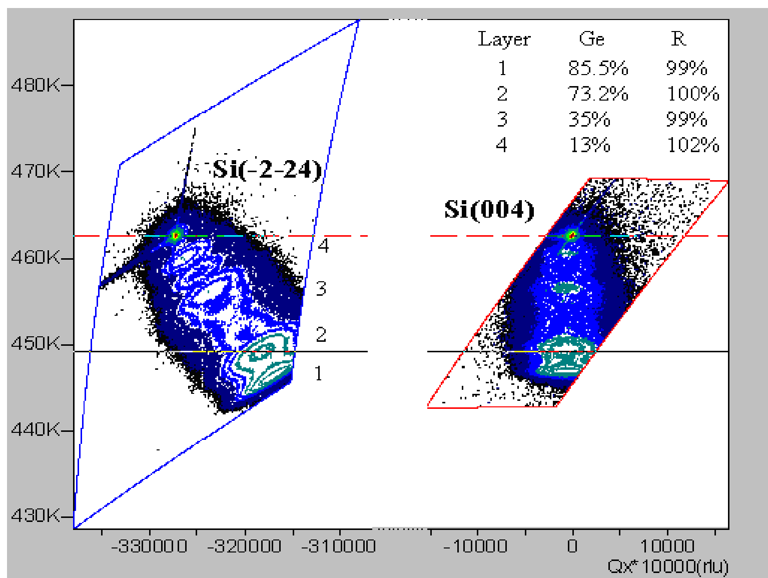



**FIGURE 3.25**
High-resolution XRD 004 and $\bar{2}\bar{2}4$ reciprocal space maps.

were analyzed with four epilayers having increasing thickness but different Ge gradient in the SiGe heterolayers. The reciprocal lattice points (RLPs) of the $Si_{1-x}Ge_x$ layers are seen to be at the bottom of the Si substrate due to their larger lattice constants and are also found to spread along the diffraction vectors because of the varying Ge composition (i.e., varying lattice constant).

### 3.4.2 SIMS

The outdiffusion of Ge from the SiGe underlayer into the thin Si cap layer in strained-Si/SiGe material systems can greatly affect the mobility of carriers in the Si channel. This outdiffusion generally takes place during epitaxial growth due to high deposition temperature and rate, as well as point defects introduced by ion implantation processes and annealing in different ambient. The Si cap layer is very thin, on the order of 10 nm, and the lattice-mismatch induced undulations (cross-hatch pattern) can be significant. Achieving the depth resolution required to obtain information on the Ge outdiffusion is very challenging. Ge interfacial abruptness is a critical aspect in the proper integration of strained-Si/SiGe into real devices.

Ultra low energy (ULE) secondary ion mass spectrometry (SIMS) are commonly used to characterize the thickness of the strained-Si cap layer, Ge composition, Ge outdiffusion (interface abruptness), and the level of atmospheric contaminants. Fig. 3.26 shows the typical composition profiles of Ge, C, and

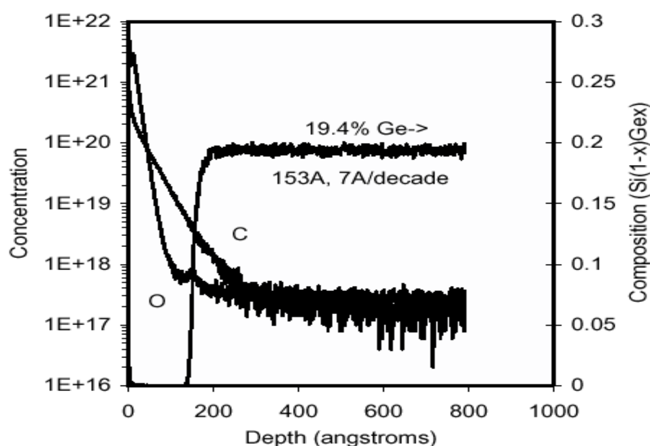



**FIGURE 3.26**
Typical SIMS profile of strained-Si on SiGe buffer layer. A single SIMS profile provides the thickness of the strained-Si cap layer, Ge composition, interface abruptness, and atmospheric contamination with high sensitivity.

O, acquired and quantified simultaneously in one ULE SIMS analysis. The Ge composition may be obtained with high accuracy up to 70% Ge and with very low detection limits for C and O.

### 3.4.3 Raman

Raman spectroscopy is commonly used to verify that the Si layer is indeed strained or not. In this study, several strained-Si films grown using GSMBE and ultralow pressure chemical vapor deposition (ULPCVD) have been used. The detail of the film growth and their properties may be found in references [60, 61]. Raman spectra were registered in backscattering geometry using micro-Raman system RENISHAW 1000 equipped with a Leica microscope. The excitation wavelength was 514.5 nm from an $Ar^+$ ion laser with a typical laser power level of $\sim$ 10 mW. In order to define the position of the phonon lines with greater accuracy, the spectral lines were fitted with a Lorentzian function.

Raman spectrum of the relaxed-SiGe layer with a thin strained-Si layer on top is shown in Fig. 3.27. The main peak at 504.8 $cm^{-1}$ belongs to Si-Si stretching mode from the SiGe layer while the shoulder at 510.8 $cm^{-1}$ belongs to the Si-Si phonon mode in the strained-Si layer. Other groups of phonon modes observed in the range of 450-370 $cm^{-1}$ and 250-325 $cm^{-1}$ are associated with the SiGe layer and belong to the Si-Ge and Ge-Ge stretching modes, respectively. The exact position of all aforementioned peaks depends on stress (for pure Si) and on stress and Ge content, x, for the SiGe layer.

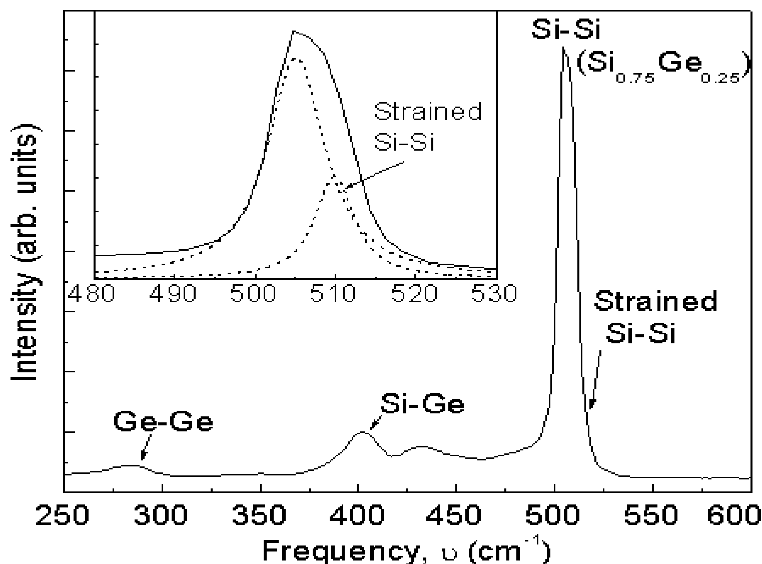



**FIGURE 3.27**
Raman spectrum of a thin strained-Si layer grown at 800$^o$C on a step-graded relaxed-SiGe buffer. After C. K. Maiti et al., *Thin Solid Films*, Vol. 462-463, 2004(8084). With permission.

Using the Si-Si and Si-Ge peak positions, it is found that the SiGe layer beneath the strained-Si layer is 100% relaxed with a Ge content of 27.2±2%.

Visible and ultraviolet Raman spectroscopy can be used effectively to map the strain related with cross-hatch pattern observed in SiGe (and in relatively thick strained-Si layers), which is usually developed in low lattice mismatch heterostructures. The Raman measurements demonstrate that the pattern is similar in both SiGe and strained-Si layers. Mapping of strain by Raman spectroscopy offers an effective nondestructive tool for controlling at the same time both the strain magnitude and its lateral fluctuation associated with cross-hatch patterns.

In Fig. 3.28 three distinct lines with strain and composition-dependent energies are found which are ascribed to Si-Si vibrations (505.7 cm$^{-1}$), localized Si-Si vibrations perturbed by neighboring Ge (431.1 cm$^{-1}$) and Si-Ge (403.6 cm$^{-1}$). Smaller peaks are due to the strained-Si cap layer and the Si substrate. The frequency shifts of the Si-Si and Si-Ge lines allow the simultaneous calculation of composition and residual strain (relaxation) in the SiGe layer using empirical relations for the composition-and strain-induced frequency shift.

Raman spectroscopy can be used to measure, at the same time, not only the Ge concentration in $Si_{1-x}Ge_x$ samples, but also the strain developed in these structures as the result of Si and Ge lattice mismatch. Three distinguished peaks at $\sim$ 500 cm$^{-1}$ (due to Si-Si vibration), $\sim$ 400 cm$^{-1}$ (due to Si-Ge

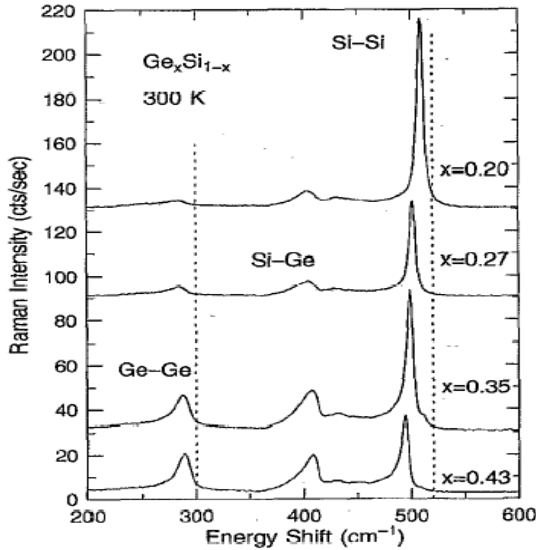



**FIGURE 3.28**
Raman spectra for epitaxial $Si_{1-x}Ge_x$ layers of various alloy composition at room temperature. After P. M. Mooney et al., *Appl. Phys. Lett.*, Vol. 62, 1993(2069-2071). With permission.

vibration) and $\sim$ 280 cm$^{-1}$ (due to Ge-Ge vibration) depend on actual x concentration of Ge and on the in plane strain. The phonon frequency $\omega$ Si-Si (in cm$^{-1}$) of Si-Si vibration in SiGe is related to x and strain $\epsilon$ by the relation [62]:

$$\omega_{Si-Si} = 520.5 - 62x - 815\epsilon \tag{3.14}$$

which means that the strain in percentage can be calculated as follows:

$$\epsilon(\%) = 0.123\Delta\Omega \tag{3.15}$$

where $\Delta\Omega$ is the shift of phonon frequency from the same peak in the fully relaxed structure (being $Si_{1-x}Ge_x$ with peak at 520.5-62x or pure Si with peak at 520.5 cm$^{-1}$). Fig. 3.29 shows the the baseline intensity, the data fit using a double Lorentzian line shape in order to the determine the intensity, Raman shift and full-width half-maximum (FWHM) of the peaks arising from the Si-Si vibrational mode in the SiGe buffer layer and the strained-Si cap layer as shown in Fig. 3.29.

Raman analysis has also been used for the characterization of process-induced strain generated by recessed SiGe S/D. Nouri et al. [63] have reported the impact of process parameters on the performance of strain-enhanced p-MOSFETs with recessed SiGe S/D. Figs. 3.30 and 3.31 show the results of

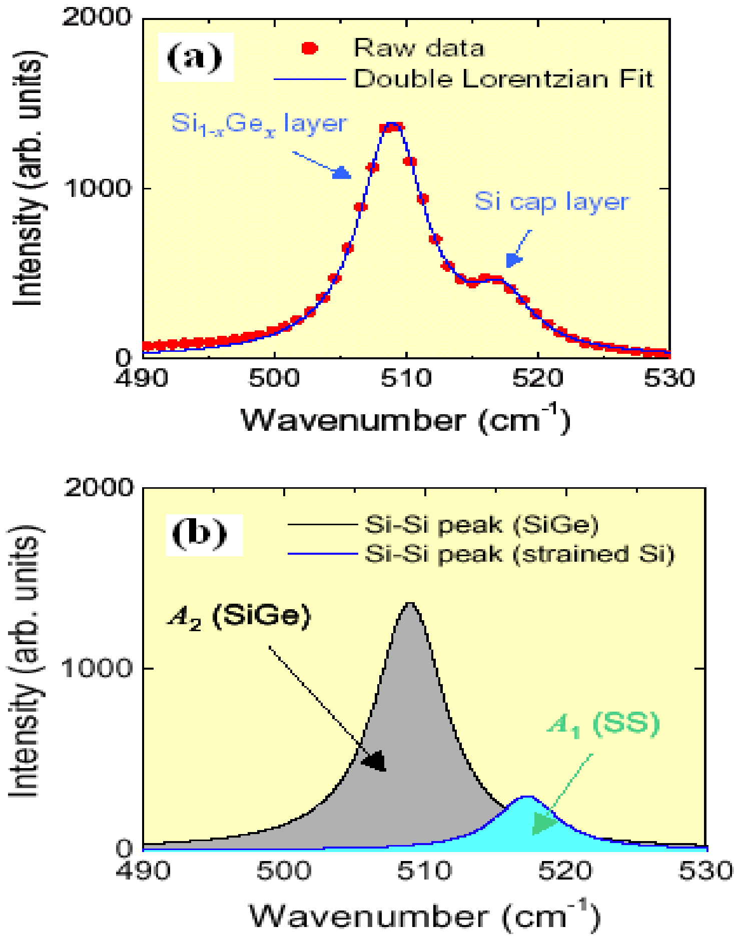



**FIGURE 3.29**

(a) Si-Si bands of SiGe (502.2 cm$^{-1}$) and of strained-Si (512.2 cm$^{-1}$) from a Raman spectrum and (b) Raman intensity map of strain strained-Si layer. After P. M. Mooney, 2003 International Conference on Characterization and Metrology for ULSI Technology, March 24-27, 2003.

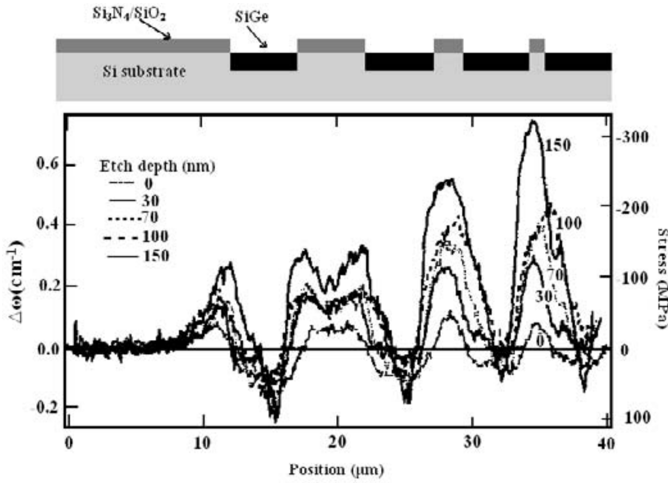



**FIGURE 3.30**
Influence of etch depth (- SiGe thickness), on stress in the channel measured by micro-Raman spectroscopy. The stress is calculated from the Raman shift assuming uniaxial stress. All samples have 20% Ge. The stress becomes more compressive with deeming channel width and increasing etch depth. The control sample (0 nm) has no SiGe. After F. Nouri et al., *IEEE IEDM Tech. Dig.*, 2004(1055-1058). With permission.

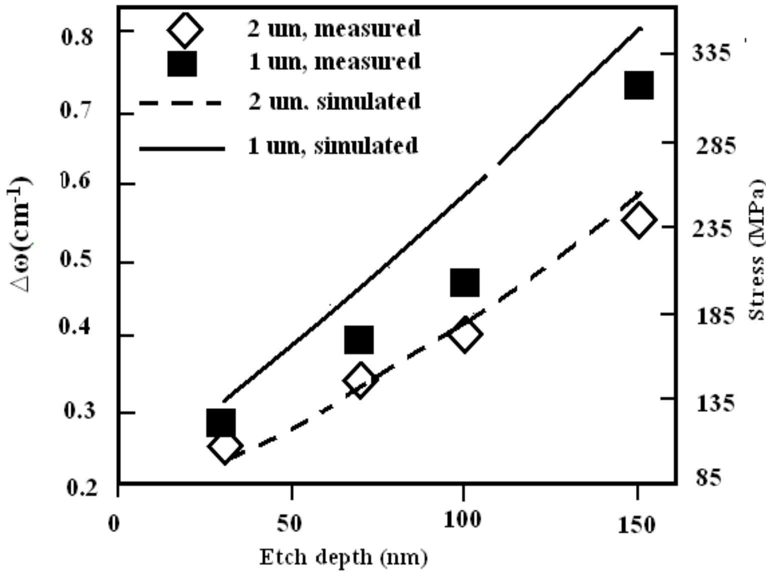



**FIGURE 3.31**
Measured vs. simulated stress as a function of etch depth for 1 and 2 $\mu$m spacings. After F. Nouri et al., *IEEE IEDM Tech. Dig.*, 2004(1055-1058). With permission.

micro-Raman spectroscopy measurements on test structures consisting of long lines (channels) of various widths, separated by 5 $\mu$m of recessed SiGe S/D. The stress induced by the SiGe in silicon (both between and under the SiGe areas) is calculated from the Raman frequency shift ($\Delta\omega$) assuming uniaxial stress: $\sigma$(MPa) = -434 $\Delta\omega$ (cm$^{-1}$). The measurements confirm that the SiGe induces compressive stress in the channel area.

## 3.5 Engineered Substrates

### 3.5.1 Mobility Comparison

Fig. 3.32 schematically illustrates various heterostructure substrates that have been used to investigate biaxial strain and high-mobility channel materials, including epitaxial structures on bulk substrates and on-insulator implementations. As discussed before, the on-insulator ("OI") substrates are generally fabricated from bulk structures by a combination of epitaxial growth, wafer bonding, and delamination or etch-back methods, which retain the strain state of the layers. For example, strained-Si directly on insulator (SSDOI), illustrated in Fig. 3.32(e), is derived from biaxial strained-Si/relaxed-$Si_{1-x}Ge_x$ by transfer of the epitaxial layers and removal of the relaxed-$Si_{1-x}Ge_x$ virtual substrate, leaving a strained-Si layer directly in contact with silicon dioxide [41, 64]. The on-insulator technologies provide a pathway to implementing mobility enhancement in partially or fully depleted devices, in ultrathin-body MOSFETs, or nonplanar (e.g., double-gate) MOSFETs.

Fig. 3.33 compares effective mobility in bulk MOSFETs for (a) electrons in strained-Si/relaxed-SiGe and (b) holes in Si-channel and strained-Si/$Si_{1-y}Ge_y$ dual-channel heterostructures. For the strained-Si channel p-MOSFETs, the data from Rim et al. is for biaxial strained-Si/relaxed $Si_{1-x}Ge_x$ (x = 28%) [65], and the data from Ghani et al. [66] is for process-induced compressively strained-Si implemented by the growth of SiGe in the source/drain regions. For dual-channel structures, the notation y/x represents the Ge mole fraction in the strained-$Si_{1-y}Ge_y$ layer and in the relaxed-$Si_{1-x}Ge_x$ substrate, respectively. The highest hole mobility is obtained for a pure Ge channel strained in biaxial compression to a relaxed-$Si_{0.5}Ge_{0.5}$ substrate [67]. Other dual-channel data is from Leitz et al. and Jung et al. [39, 68]. The curves marked "Universal" are for unstrained-Si, from Takagi et al. [69] are also shown for comparison.

Dual-channel heterostructures (Fig. 3.32(b)) use a combination of strained-Si and strained-$Si_{1-y}Ge_y$ to enable simultaneously high electron and hole mobilities [39, 67, 68]. In addition, because of the high Ge content and compressive strain, these structures offer significantly higher hole mobility than either biaxial tensile strained-Si (Fig. 3.33(b)) or process-induced strained-Si

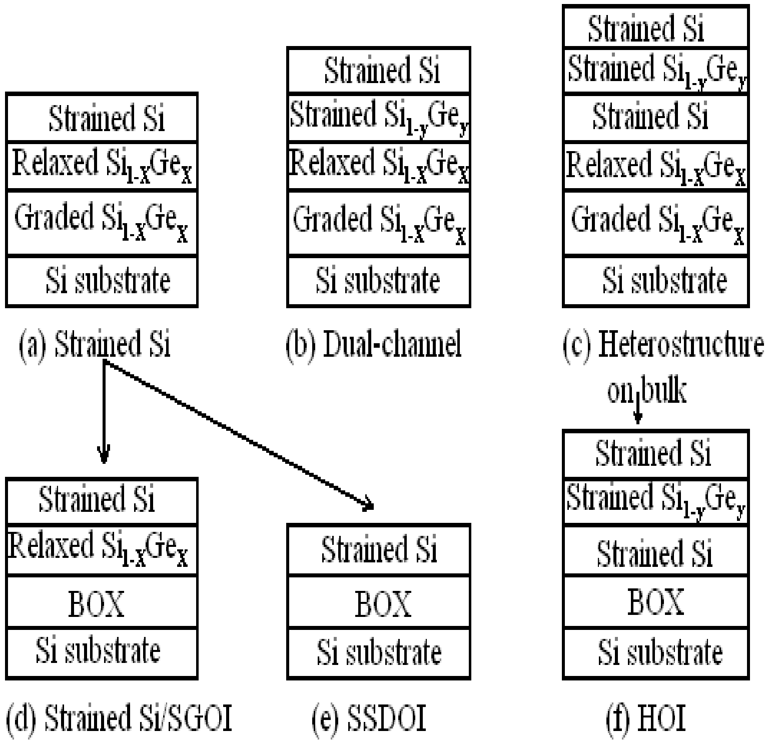



**FIGURE 3.32**
Schematics of various reported heterostructure substrates used to enhanced mobility. (a) biaxial strained-Si on relaxed-SiGe, (b) dual-channel structure (y > x), and (c) heterostructure on bulk; and on-insulator implementations on buried oxide (BOX) layers such as (d) strained-Si/relaxed-$Si_{1-x}Ge_x$ on insulator (SGOI), (e) strained-Si directly on insulator (SSDOI), and (f) heterostructure on insulator (HOI). After J. L. Hoyt et al., Enhanced Mobility CMOS, Lecture Materials.

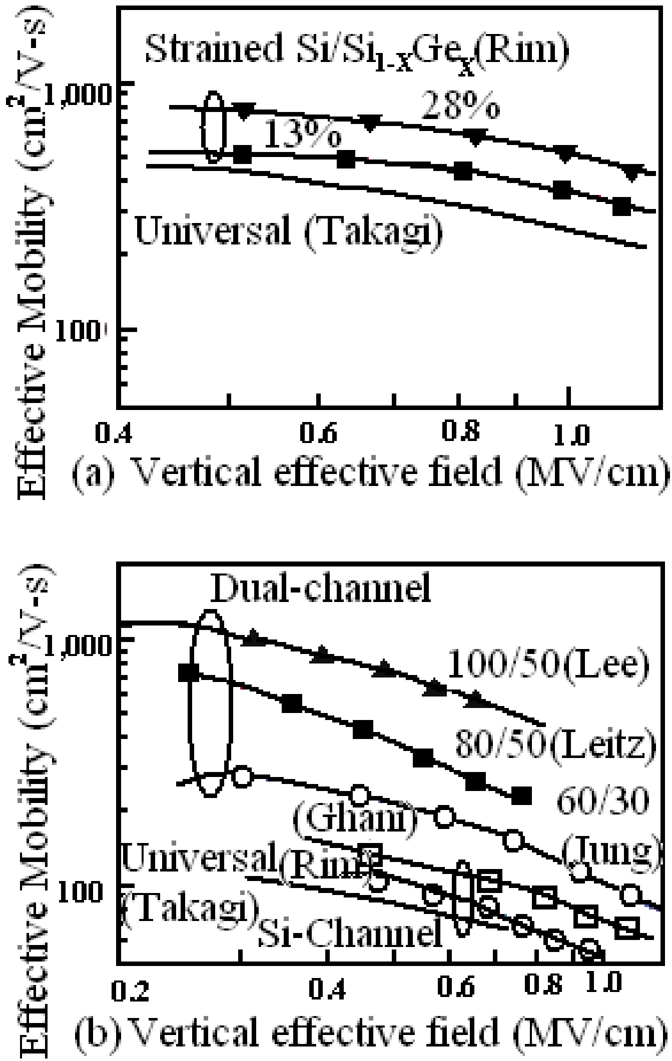



**FIGURE 3.33**

Comparison of measured effective mobility in bulk technologies for (a) electrons in strained-Si/relaxed-$Si_{1-x}Ge_x$, and (b) holes in various Si-channel and strained-Si/strained-SiGe dual-channel implementations. See text for detail. After J. L. Hoyt et al., Enhanced Mobility CMOS, Lecture Materials.

(Fig. 3.33(b)). As illustrated in Fig. 3.33(b), the hole mobility increases with increasing Ge fraction in the strained-$Si_{1-y}Ge_y$ channel. A hole-mobility enhancement factor of approximately 10×, relative to unstrained-Si, is obtained for a p-MOSFET with a thin strained-Si layer (∼ 5 nm) on top of a strained Ge channel (∼ 12 nm) on a relaxed-$Si_{0.5}Ge_{0.5}$ virtual substrate [67].

The electron mobility in the strained-Si channel is enhanced by approximately a factor of 2 for the same structure. The high hole mobility results from a combination of the small Ge in-plane hole effective mass, the biaxial compression-induced strain splitting of the valence bands, and the use of a thin Si cap layer, which enables a low interface state density to be obtained at the semiconductor/insulator interface. In Ge surface-channel MOSFETs, where the gate dielectric is formed directly on Ge without an intermediate Si layer, reported hole mobility enhancements are in the range of 1.4 to 2× [70, 71, 72], and n-MOSFET performance is disappointing [73]. The use of Si-compatible gate insulator technology makes the strained-Si/strained-Ge dual-channel heterostructure very attractive.

### 3.5.2 Thermal Conductivity

The success of miniaturization in electronics has owed largely to how conveniently CMOS transistors scale down. Shrinking the transistor primarily entails shortening its gate length and results in higher operating speed (thus higher operating current) leading to higher temperature rise with each new technology generation. It is thus important to address materials solutions that enhance the thermal conductivity of the substrates, thereby mitigating potential hot spots and the overall overheating of circuit components. Materials with higher thermal conductivity than Si in the immediate vicinity of the transistors could potentially keep circuits cooler due to their more efficient heat removal. For example, single crystal diamond and single crystal silicon carbide (SiC) have thermal conductivities, 20 W $cm^{-1}K^{-1}$ and 3-4 $cm^{-1}K^{-1}$, respectively, that are much greater than that of Si, 1.5 W $cm^{-1}K^{-1}$. Thus, it would be favorable to have diamond and SiC regions incorporated into chip architectures for thermal management reasons.

To reduce the operating temperature of the Si transistors, the transistor regions need to be placed on top of either (a) a handle wafer consisting of higher thermal conductivity material where heat can be dissipated more efficiently to a heat sink and (b) a high thermal conductivity layer that itself resides on top of conventional Si substrate, where heat removal from the active circuit is not changed much, but heat can be spread more uniformly across the entire area, thus reducing the potential for hot spots. In SOI structures the insulator is normally a thin layer of thermally grown $SiO_2$. A high-quality dielectric layer with better thermal properties than $SiO_2$ would be of interest for improving heat dissipation in SOI circuits. Amorphous, polycrystalline or crystalline aluminum oxide, having thermal conductivity 10-30 times higher than $SiO_2$, may be used to improve heat dissipation during SOI circuit operation.

Toward thermal management problems, the use of isotopically pure Si could be a possible solution. The natural abundance of Si isotopes is 92.17% 28 Si, 4.71% 29 Si and 3.12% 30 Si. The density (concentration) of imperfections in the form of isotopes (8%) far exceeds any other defects (e.g., doping) and is the dominant scattering mechanism for the phonons that transport heat. The presence of several percent of the more massive 29 Si and 30 Si isotopes in a predominantly 28 Si crystal acts to scatter the phonons (i.e., lattice vibrations) that transport most of the heat through the crystal, decreasing its thermal conductivity. If the crystal contains isotopes of one mass only, e.g., 28 Si, this component of phonon scattering is eliminated, increasing the thermal conductivity, in principle. It has been shown experimentally and theoretically that isotopic purification of group-IV elements (C, Si, Ge) leads to significant increases in thermal conductivity. A 5-10% reduction in overall thermal resistance is consistently measured.

In the case of isotopically pure silicon (99.9% Si) a peak thermal conductivity of 30,000 W/m-K is measured at 20 K [74]. This represents an six-fold improvement over natural silicon. However, there is significant scatter in the data for the extent of the improvement (10% to 60%) in the 300-400 K range. For "undoped" silicon at room temperature the range of measured thermal conductivity is: 142-148 W/m-K (natural Si) vs. 165-227 W/m-K Si. Thermal conductivity degrades with increased doping by about 20% for both natural Si and Si for an acceptor concentration of $2\times10^{19}$ cm$^{-3}$. A novel layer transfer technique deploying isotopically pure Si has recently been reported by Kizilyalli et al. [74]. In case of RF LDMOS transistors fabricated using this methodology, a reduction of about 5-7$^o$C in transistor average temperature is measured by infrared microscopy as compared with devices fabricated using natural Si. Thus, it is likely that the use of $^{28}$Si may complement other strategies for heat extraction from integrated circuits.

However, this approach is prohibitively expensive today due to the high cost of isotope separation and the large amount of isotopically pure Si required for a large diameter crystal. Alternatively, a thin layer of isotopically pure Si can be deposited as the layer of an epitaxial wafer or SOI wafer. The cost of depositing a thick $^{28}$Si epitaxial layer in a conventional 200 mm single wafer reactor is estimated at $400, but may be substantially reduced by leveraging high-volume production methodologies.

A significant drawback to the SSOI, SGOI, GOI, and strained-Si on relaxed-SiGe platforms mentioned above for high-power devices is their dramatically reduced thermal conductivity values near the active device regions relative to bulk silicon substrate. Though its presence drastically reduces parasitic capacitances, an intermediate oxide layer has a thermal conductivity value of 0.014 W cm$^{-1}$ K$^{-1}$ at 300 K, nearly 2 orders of magnitude lower than bulk-Si, which serves to significantly reduce heat extraction from the device channel.

## 3.6 Summary

In this chapter, we reviewed the recent progress made in new engineered substrate materials and their potential in microelectronics with a view in mind that: (a) new materials needed to be incorporated to continue downscaling and (b) the performance advantage available from new materials. Improving device performance through carrier mobility enhancement is of great importance for CMOS scaling. It can be done at the device level by introducing a local stressing layer (e.g., strained-SiGe or silicon nitride film) within the transistor structure and at the wafer level either through the introduction of stress in the active layer or by optimizing crystal orientation for both pMOS and nMOS devices.

With the availability of SiGe-free strained-Si layers on SOI substrates (SSOI) that promise to break through today's looming barriers of power dissipation - an increasingly critical roadblock to higher-performing chips will be removed. Since wafer-level (global) strain is not dependent on IC design, SSOI substrates will enable a wider range of high-speed, low-power IC applications, including those with high-performance logic cores. However, some challenges for engineered substrates need to be overcome, such as the lower thermal conductivity of most engineered substrates such as strained-Si on SiGe or SOI platform.

Strained layer and epitaxy process characterization is metrology intensive. Final strained-Si quality is strongly dependent on earlier growth stages. CMP complicates the logistics and adds significantly to the cost. Commercialization is focusing on higher throughput CVD epitaxy processes. New engineered substrate materials, like strained-Si, will aid the extension of Moore's law and strained-Si can be considered as the beginning of a new roadmap on engineered substrates.

Manufacturing practices in the CMOS industry are currently experiencing a discontinuity, where new approaches are necessary in order to maintain improved product capability within reasonable economic constraints. Strained-Si technology is a very promising technology for CMOS manufacturers and circuit designers. Strained-Si has the potential to provide great performance benefits with minimal disruption to standard CMOS manufacturing infrastructure.

## References

[1] H. Kroemer, "Theory of a wide-gap emitter for transistors," *Proc. IRE*,

vol. 45, pp. 1535–1537, 1957.

[2] F. C. Frank and J. H. van der Merwe, "One-dimensional dislocations. II. misfitting monolayers and oriented overgrowth," *Proc. Roy. Soc.*, vol. A198, pp. 216–225, 1949.

[3] W. Shockley, US Patent No. 2569347, 1951.

[4] E. A. Fitzgerald, Y. H. Xie, D. Monroe, P. J. Silverman, J. M. Kuo, and A. R. Kortan, "Relaxed $Ge_xSi_{1-x}$ structures for III-V integration with Si and high mobility two-dimensional electron gases in Si," *J. Vac. Sci. Technol. B*, vol. 10, pp. 1807–1819, 1992.

[5] D. J. Paul, "Si/SiGe heterostructures: from material and physics to devices and circuits," *Semicond. Sci. Technol.*, vol. 19, pp. R75–R108, 2004.

[6] E. Kasper, H. J. Herzog, and H. Kibbel, "A one dimensional SiGe superlattice grown by UHV epitaxy," *Appl. Phys.*, vol. 8, pp. 199–205, 1975.

[7] J. C. Bean, L. C. Feldman, A. T. Fiory, S. Nakahara, and I. K. Robinson, "$Ge_xSi_{1-x}$/Si strained layer superlattice growth by molecular beam epitaxy," *J. Vac. Sci. Technol. A*, vol. 2, pp. 436–440, 1984.

[8] J. Matthews, S. Mader, and T. Light, "Accommodation of misfit across the interface between crystals of semiconducting elements or compounds," *J. Appl. Phys.*, vol. 41, pp. 3800–3804, 1970.

[9] D. C. Houghton, "Strain relaxation kinetics in $Si_{1-x}Ge_x$/Si heterostructures," *J. Appl. Phys.*, vol. 70, pp. 2136–2151, 1991.

[10] B. W. Dodson and J. Y. Tsao, "Relaxation of strained-layer semiconductor structures via plastic flow," *Appl. Phys. Lett.*, vol. 51, pp. 1325–1327, 1987.

[11] J. H. van der Merwe, "Crystal interfaces. Part II. finite overgrowths," *J. Appl. Phys.*, vol. 34, pp. 123–127 (see also Erratum p. 3420, (1963)), 1963.

[12] J. W. Matthews and A. E. Blakeslee, "Defects in epitaxial multilayers - III. preparation of almost perfect multilayers," *J. Crystal Growth*, vol. 32, pp. 265–273, 1976.

[13] K. Eberl, S. S. Iyer, S. Zollner, J. C. Tsang, and F. K. LeGoues, "Growth and strain compensation effects in the ternary $Si_{1-x-y}Ge_xC_y$ alloy system," *Appl. Phys. Lett.*, vol. 60, pp. 3033–3035, 1992.

[14] A. St. Amour, L. D. Lanzerotti, C. L. Chang, and J. C. Sturm, "Optical and electrical properties of $Si_{1-x-y}Ge_xC_y$ thin films and devices," *Thin Solid Films*, vol. 294, pp. 112–117, 1997.

[15] C. K. Maiti, L. K. Bera, and S. Chattopadhyay, "Strained-Si heterostructure field effect transistors," *Semicond. Sci. Technol.*, vol. 13, pp. 1225–1246, 1998.

[16] M. L. Lee, E. A. Fitzgerald, M. T. Bulsara, M. T. Currie, and A. Lochtefeld, "Strained Si, SiGe, and Ge channels for high-mobility metal-oxide-semiconductor field-effect transistors," *J. Appl. Phys.*, vol. 97, pp. 011101–1–011101–27, 2005.

[17] D. J. Paul, "Silicon-germanium strained layer materials in microelectronics," *Advanced Materials*, vol. 11, pp. 191–204, 1999.

[18] U. Konig and H. Daembkes, "SiGe HBTs and HFETs," *Solid-State Electron.*, vol. 38, pp. 1595–1602, 1995.

[19] R. People, "Physics and applications of $Ge_xSi_{1-x}$ strained-layer heterostructures," *IEEE J. Quantum Elec.*, vol. 22, pp. 1696–1710, 1986.

[20] D. K. Nayak and S. K. Chun, "Low-field hole mobility of strained-Si on (100) $Si_{1-x}Ge_x$ substrate," *Appl. Phys. Lett.*, vol. 64, pp. 2514–2516, 1994.

[21] S. C. Jain, H. J. Osten, B. Dietrich, and H. Rucker, "Growth and properties of strained-$Si_{1-x-y}Ge_xC_y$ layers," *Semicond. Sci. Technol.*, vol. 10, pp. 1289–1302, 1995.

[22] H. Jorke and H. J. Herzog, "Mobility enhancement in modulation doped Si-$Si_{1-x}Ge_x$ superlattice grown by molecular beam epitaxy," *Proc. Silicon Molecular Beam Epitaxy, Electro. Chem. Soc.*, p. 352, 1985.

[23] G. Abstreiter, T. Wolf, H. Jorke, and H. J. Herzog, "Strain-induced two-dimensional electron gas in selectively doped Si/$Si_xGe_{1-x}$ superlattice," *Phys. Rev. Lett.*, vol. 54, pp. 2441–2445, 1985.

[24] F. K. LeGoues, M. M. Mooney, and J. O. Chu, "Crystallographic tilting resulting from nucleation limited relaxation," *Appl. Phys. Lett.*, vol. 62, pp. 140–142, 1993.

[25] F. K. LeGoues, B. S. Meyerson, J. F. Morar, and P. D. Kirchner, "Mechanism and conditions for anomalous strain relaxation in graded thin films and superlattices," *J. Appl. Phys.*, vol. 71, pp. 4230–4243, 1992.

[26] P. M. Mooney, J. L. Jordan-Sweet, K. Ismail, J. O. Chu, R. M. Feenstra, and F. K. LeGoues, "Relaxed $Si_{0.7}Ge_{0.3}$ buffer layers for high-mobility devices," *Appl. Phys. Lett.*, vol. 67, pp. 2373–2375, 1995.

[27] J. W. P. Hsu, E. A. Fitzgerald, Y. H. Xie, P. J. Silverman, and M. J. Cardillo, "Surface morphology of relaxed $Ge_xSi_{1-x}$ films," *Appl. Phys. Lett.*, vol. 61, pp. 1293–1295, 1992.

[28] M. A. Lutz, R. M. Feenstra, F. K. LeGoues, P. M. Mooney, and J. O. Chu, “Influence of misfit dislocations on the surface morphology of $Si_{1-x}Ge_x$ films,” *Appl. Phys. Lett.*, vol. 66, pp. 724–726, 1995.

[29] E. A. Fitzgerald, Y.-H. Xie, M. L. Green, D. Brasen, A. R. Kortan, J. Michel, Y.-J. Mii, and B. E. Weir, “Totally relaxed $Ge_xSi_{1-x}$ layers with low threading dislocation densities grown on Si substrates,” *Appl. Phys. Lett.*, vol. 59, pp. 811–813, 1991.

[30] Y.-H. Xie, E. A. Fitzgerald, D. Monroe, G. P. Watson, and P. J. Silverman, “From relaxed GeSi buffers to field effect transistors: current status and future prospects,” *Jap. J. Appl. Phys.*, vol. 33, pp. 2372–2377, 1994.

[31] K. Rim, J. L. Hoyt, and J. F. Gibbons, “Fabrication and analysis of deep submicron strained-Si N-MOSFETs,” *IEEE Trans. Electron Dev.*, vol. 47, pp. 1406–1415, 2000.

[32] R. M. Sieg, J. A. Carlin, S. A. Ringel, M. T. Currie, S. M. Ting, T. A. Langdo, G. Taraschi, E. A. Fitzgerald, and B. M. Keyes, “High minority-carrier lifetimes in GaAs grown on low-defect-density Ge/GeSi/Si substrate,” *Appl. Phys. Lett.*, vol. 73, pp. 3111–3113, 1998.

[33] D. K. Nayak, N. Usami, H. Sunamura, S. Fukatsu, and Y. Shiraki, “Band-edge photoluminescence of SiGe/strained-Si/SiGe type-II quantum wells on Si(100),” *Jap. J. Appl. Phys.*, vol. 32, pp. L1391–L1393, 1993.

[34] D. K. Nayak, N. Usami, S. Fukatsu, and Y. Shiraki, “Band-edge photoluminescence of SiGe/strained-Si/SiGe type-II quantum wells on Si(100),” *Appl. Phys. Lett.*, vol. 63, pp. 3509–3511, 1993.

[35] S. Fukatsu, H. Yoshida, A. Fujiwara, Y. Takahashi, and Y. Shiraki, “Spectral blue shift of photoluminescence in strained layer $Si_{1-x}Ge_x$/Si quantum well structures grown by gas source Si MBE,” *Appl. Phys. Lett.*, vol. 61, pp. 804–806, 1992.

[36] Y. Kato, S. Fukatshu, and Y. Shiraki, “Postgrowth of a silicon contact layer on air-exposed $Si_{1-x}Ge_x$/Si single quantum well grown by gas source molecular beam epitaxy for use in an electroluminescent device,” *J. Vac. Sci. Technol.*, vol. B13, pp. 111–117, 1995.

[37] Y.-H. Xie, D. Monroe, E. A. Fitzgerald, P. J. Silverman, F. A. Thiel, and G. P. Watson, “Very high mobility two-dimensional hole gas in Si/GeSi/Ge structures grown by molecular beam epitaxy,” *Appl. Phys. Lett.*, vol. 63, pp. 2263–2265, 1993.

[38] G. Hock, E. Kohn, C. Rosenblad, H. von Kanel, H.-J. Herzog, and U. Konig, “High hole mobility in $Si_{0.17}Ge_{0.83}$ channel metal-oxide-

semiconductor field-effect transistors grown by plasma-enhanced chemical vapor deposition," *Appl. Phys. Lett.*, vol. 76, pp. 3920–3922, 2000.

[39] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, "Hole mobility enhancements in strained-Si/$Si_{1-y}Ge_y$ p-type metal-oxide-semiconductor field-effect transistors grown on relaxed $Si_{1-x}Ge_x$ ($x < y$) virtual substrates," *Appl. Phys. Lett.*, vol. 79, pp. 4246–4248, 2001.

[40] M. L. Lee and E. A. Fitzgerald, "Strained Si/strained Ge dual-channel heterostructures on relaxed $Si_{0.5}Ge_{0.5}$ for symmetric mobility p-type and n-type metal-oxide-semiconductor field-effect transistors," *Appl. Phys. Lett.*, vol. 83, pp. 4202–4204, 2003.

[41] T. A. Langdo, A. Lochtefeld, M. T. Currie, R. Hammond, V. K. Yang, J. A. Carlin, C. J. Vineis, G. Braithwaite, H. Badawi, M. T. Bulsara, and E. A. Fitzgerald, "Preparation of novel SiGe-free strained-Si on insulator substrates," *Proc. IEEE Int. SOI Conf.*, pp. 211–212, 2002.

[42] S. Fukatsu, Y. Ishikawa, and N. Shibata, "SiGe-based semiconductor-on-insulator substrate created by low-energy separation-by-implanted-oxygen," *Appl. Phys. Lett.*, vol. 72, pp. 3485–3487, 1998.

[43] G. Taraschi, T. A. Langdo, M. T. Currie, E. A. Fitzgerald, and D. A. Antoniadis, "Relaxed SiGe-on-insulator fabricated via wafer bonding and etch back," *J. Vac. Sci. Technol. B*, vol. 20, pp. 725–727, 2002.

[44] L. J. Huang, J. O. Chu, D. F. Canaperi, C. P. D'Emic, R. M. Anderson, S. J. Koester, and H.-S. P. Wong, "SiGe-on-insulator prepared by wafer bonding and layer transfer for high-performance field-effect transistors," *Appl. Phys. Lett.*, vol. 78, pp. 1267–1269, 2001.

[45] T. Mizuno, S. Takagi, N. Sugiyama, H. Satake, A. Kurobe, and A. Toriumi, "Electron and hole mobility enhancement in strained-Si MOSFETs on SiGe-on-insulator substrates fabricated by SIMOX technology," *IEEE Electron Dev. Lett.*, vol. 21, pp. 230–232, 2000.

[46] T. Tezuka, N. Sugiyama, and S. Takagi, "Fabrication of strained-Si on an ultrathin SiGe-on-insulator virtual substrate with a high-Ge fraction," *Appl. Phys. Lett.*, vol. 79, pp. 1798–1800, 2001.

[47] S. Nakaharai, T. Tezuka, N. Sugiyama, Y. Moriyama, and S. Takagi, "Characterization of 7 nm thick strained Ge-on-insulator layer fabricated by Ge-condensation technique," *Appl. Phys. Lett.*, vol. 83, pp. 3516–3518, 2003.

[48] T. S. Drake, C. NyChleirigh, M. L. Lee, A. J. Pitera, E. A. Fitzgerald, D. A. Antoniadis, D. H. Anjum, J. Li, R. Hull, N. Klymko, and J. L. Hoyt, "Effect of rapid thermal annealing on strain in ultrathin strained silicon on insulator layers," *Appl. Phys. Lett.*, pp. 875–877, 2003.

[49] I. Aberg, O. O. Olubuyide, C. N. Chleirigh, I. Lauer, D. A. Antoniadis, J. Li, R. Hull, and J. L. Hoyt, "Electron and hole mobility enhancements in sub-10 nm thick strained silicon directly on insulator fabricated by a bond and etch-back technique," *IEEE VLSI Tech. Symp. Dig.*, pp. 52–53, 2004.

[50] S. H. Christiansen, R. Singh, I. Radu, M. Reiche, U. Gosele, D. Webb, S. Bukalo, and B. Dietrich, "Strained silicon on insulator (SSOI) by wafer bonding," *Mat. Sci. Semicond. Proc.*, vol. 8, pp. 197–202, 2005.

[51] Soitec Inc., http://www.soitec.com/, 2005.

[52] H. Yin, K. D. Hobart, R. L. Peterson, F. J. Kub, S. R. Shieh, T. S. Duffy, and J. C. Sturm, "Fully-depleted strained-Si on insulator NMOSFETs without relaxed SiGe buffers," *IEEE IEDM Tech. Dig.*, pp. 53–56, 2003.

[53] S. Takagi, J. L. Hoyt, J. J. Welser, and J. F. Gibbons, "Comparative study of phonon-limited mobility to two-dimensional electrons in strained and unstrained Si metal-oxide-semiconductor field-effect transistors," *J. Appl. Phys.*, vol. 80, pp. 1567–1577, 1996.

[54] T. Tezuka, N. Sugiyama, T. Mizuno, M. Suzuki, and S. Takagi, "A novel fabrication technique of ultrathin and relaxed SiGe buffer layers with high Ge fraction for the sub-100 nm strained silicon-on-insulator MOSFETs," *Jpn. J. Appl. Phys.*, vol. 40, pp. 2866–2874, 2001.

[55] T. Tezuka, N. Sugiyama, S. Takagi, and T. Kawakubo, "Dislocation-free formation of relaxed SiGe-on-insulator layers," *Appl. Phys. Lett.*, vol. 80, pp. 3560–3562, 2002.

[56] G. Taraschi, A. J. Pitera, L. M. McGill, Z.-Y. Cheng, M. L. Lee, T. A. Langdo, and E. A. Fitzgerald, "Ultrathin strained Si-on-insulator and SiGe-on-insulator created using low temperature wafer bonding and metastable stop layers," *J. Electrochem. Soc.*, vol. 151, pp. G47–G56, 2004.

[57] T. Tezuka, N. Sugiyama, T. Mizuno, and S. Takagi, "Novel fully depleted SiGe-on-insulator p-MOSFETs with high-mobility SiGe surface channels," *IEEE IEDM Tech. Dig.*, pp. 946–948, 2001.

[58] A. J. Pitera, G. Taraschi, M. L. Lee, C. W. Leitz, Z. Y. Cheng, and E. A. Fitzgerald, "Coplanar integration of lattice-mismatched semiconductors with silicon by wafer bonding Ge/$Si_{1-x}Ge_x$/Si virtual substrates," *J. Electrochem. Soc.*, vol. 151, pp. G443–G447, 2004.

[59] J. Welser, J. L. Hoyt, and J. F. Gibbons, "Growth and processing of relaxed-$Si_{1-x}Ge_x$/strained-Si structures for metal oxide semiconductor applications," *Jap. J. Appl. Phys.*, vol. 33, pp. 2419–2422, 1994.

[60] D. K. Nayak, K. Goto, A. Yutani, J. Murota, and Y. Shiraki, "High-mobility strained-Si PMOSFETs," *IEEE Trans. Electron Dev.*, vol. 43, pp. 1709–1715, 1996.

[61] C. S. Tan, W. K. Choi, L. K. Bera, K. L. Pey, D. A. Antoniadis, E. A. Fitzgerald, M. T. Currie, and C. K. Maiti, "$N_2O$ oxidation of strained-Si/relaxed-SiGe heterostructure grown by UHVCVD," *Solid-State Electron.*, vol. 45, pp. 1945–1949, 2001.

[62] K. Savano, S. Koh, Y. Shiraki, N. Usami, and K Nakagawa, "In-plane strain fluctuation in strained-Si/SiGe heterostructures," *Appl. Phys. Lett.*, vol. 83, pp. 4339–4341, 2003.

[63] F. Nouri, P. Verheyen, L. Washington, V. Moroz, I. De Wolf, M. Kawaguchi, S. Biesemans, R. Schreutelkamp, Y. Kim, M. Shen, X. Xu, R. Rooyackers, M. Jurczak, G. Eneman, K. De Meyer, L. Smith, D. Pramanik, H. Forstner, S. Thirupapuliyur, and G. S. Higashi, "A systematic study of trade-offs in engineering a locally strained pMOSFET," *IEEE IEDM Tech. Dig.*, pp. 1055–1058, 2004.

[64] K. Rim, K. Chan, L. Shi, D. Boyd, J. Ott, N. Klymko, F. Cardone, L. Tai, S. Koester, M. Cobb, D. Canaperi, B. To, E. Duch, I. Babich, R. Carruthers, P. Saunders, G. Walker, Y. Zhang, M. Steen, and M. Ieong, "Fabrication and mobility characteristics of ultrathin strained-Si directly on insulator (SSDOI) MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 47–52, 2003.

[65] K. Rim, J. Chu, H. Chen, K. A. Jenkins, T. Kanarsky, K. Lee, A. Mocuta, H. Zhu, R. Roy, J. Newbury, J. Ott, K. Petrarca, P. Mooney, D. Lacey, S. Koester, K. Chan, D. Boyd, M. Leong, and H.-S. Wong, "Characteristics and device design of sub-100 nm strained-Si n- and p-MOSFETs," *IEEE VLSI Tech. Symp. Dig.*, pp. 98–99, 2002.

[66] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson, and M. Bohr, "A 90 nm high volume manufacturing logic technology featuring novel 45 nm gate length strained silicon CMOS," *IEEE IEDM Tech. Dig.*, pp. 978–980, 2003.

[67] M. L. Lee and E. A. Fitzgerald, "Optimized strained-Si/strained-Ge dual-channel heterostructures for high mobility P- and N-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 429–432, 2003.

[68] J.-W. Jung, M. L. Lee, S. F. Yu, E. A. Fitzgerald, and D. A. Antoniadis, "Implementation of both high-hole and electron mobility in strained Si/strained-$Si_{1-y}Ge_y$ on relaxed $Si_{1-x}Ge_x$ virtual substrate," *IEEE Electron Dev. Lett.*, vol. 24, pp. 460–462, 2003.

[69] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the universality of inversion layer mobility in Si MOSFETs: Part I - Effects of substrate impurity concentration," *IEEE Trans. Electron Dev.*, vol. 41, pp. 2357–2362, 1994.

[70] H. Shang, H. Okorn-Schmidt, K. K. Chan, M. Copel, J. Ott, P. M. Kozlowski, S. Steen, R. M. Sicina, E. Sikorski, S. Cordes, H.-S. P. Wong, E. C. Jones, and W. Haensch, "High mobility p-channel germanium MOSFETs with a thin Ge oxynitride gate dielectric," *IEEE IEDM Tech. Dig.*, pp. 441–444, 2002.

[71] A. Ritenour, S. Yu, M. L. Lee, N. Lu, W. Bai, A. Pitera, E. A. Fitzgerald, D. L. Kwong, and D. A. Antoniadis, "Epitaxial strained germanium p-MOSFETs with $HfO_2$ gate dielectric and TaN gate electrode," *IEEE IEDM Tech. Dig.*, pp. 433–436, 2003.

[72] C. O. Chui, H. Kim, D. Chi, B. B. Triplett, P. C. McIntyre, and K. C. Saraswat, "A sub-400$^o$C germanium MOSFET technology with high-$\kappa$ dielectric and metal gate," *IEEE IEDM Tech. Dig.*, pp. 437–440, 2002.

[73] C. O. Chui, H. Kim, P. C. McIntyre, and K. C. Saraswat, "A germanium NMOSFET process integrating metal gate and improved high-k dielectrics," *IEEE IEDM Tech. Dig.*, pp. 437–440, 2003.

[74] I. C. Kizilyalli, H. Safar, J. Herbsommer, S. J. Burden, and P. L. Gammel, "Power transistors fabricated using isotopically purified silicon ($^{28}$Si)," *IEEE Electron Dev. Lett.*, vol. 26, pp. 404–406, 2005.

# 4

# *Electronic Properties of Engineered Substrates*

Toward strained-Si engineering, during the last 15 years major focus has been on tensilely strained (biaxial) substrates as it is possible to improve both n- and p-MOSFETs simultaneously. Although this is true, it occurs only at low electric field and high stress [1]. Currently, uniaxial stress has become the preferred method to strain the Si lattice and has been implemented into several 90-nm logic technologies [2, 3, 4]. Strain influences electrons and holes in CMOS transistors differently. Tensile strain, in which the interatomic distances in the silicon crystal are stretched, typically increases the mobility of electrons, making n-MOSFETs faster. But tensile strain may not benefit p-MOSFETs as much, and it may even slow them down. Compressive strain, in which those interatomic distances are shortened, produces essentially the opposite effects.

Uniaxial stress engineering has become a key technology to improve CMOS performance. However, the physical mechanisms of electron and hole mobility enhancement by uniaxial stress have not yet been fully understood. It is necessary to understand the physical mechanisms for modeling the electron and hole mobility. Also, the limits of uniaxial stress engineering, in terms of mobility enhancement, and its effectiveness in ballistic regime needs to be identified. It has been reported that the hole mobility enhancement obtained either via biaxial or uniaxial strain is due to the conduction band splitting and mobility enhancement with a tensile nitride-capping layer [5, 6] is similar to the biaxial strained-Si on relaxed-$Si_{1-x}Ge_x$.

Unlike the traditional approach where biaxial strain is applied into the channel from the bottom using strained-Si on relaxed-$Si_{1-x}Ge_x$ [7, 8, 9, 10], uniaxial strain is introduced from the side using $Si_{1-x}Ge_x$ in the source/drain region (for p-MOSFET) and a nitride-capping film with a large tensile stress (for n-MOSFET) [11]. While biaxial tensile stress has a great technological importance to CMOS logic technologies since it introduces advantageous strain for both n- and p-MOSFETs, it has not been introduced in commercial CMOS fabrication of device and circuits due to several integration challenges, process complexity, and cost. Furthermore, most p-MOSFETs fabricated using biaxially strained-Si, demonstrates a small hole mobility improvement at high vertical fields where the most commercial MOSFETs operate [9].

In spite of difficulties to keep short-channel effects under control and to

maintain an acceptable level of performance, advanced bulk MOSFETs appear as the most probable candidates for CMOS devices integration till 45-nm ($L_{ch}$ = 18 nm) technology node. However, a reliable and physically based model for their electrical performances is still missing. Currently available compact models rely on the usual concept of drift diffusion and saturation velocity, implicitly assuming the occurrence of a large number of scattering events in carrier transport from source to drain. However, as the carrier mean free path becomes comparable with the channel length, these assumptions are no longer valid as the devices may operate close to their ballistic limit.

In such devices the performances are no longer limited by transport mechanisms along the channel but rather by the carrier injection mechanism and scattering processes occurring close to the source end [12, 13]. These phenomena are strongly influenced by the properties of the inversion layer in this area, and in particular by its quantum 2D carrier gas nature [14], especially in highly scaled bulk devices with high channel doping and thin gate oxide. Consequently, quantization must be carefully taken into account in an improved approach for quasiballistic compact modeling. Different analytical models [12, 15] have been proposed in the literature but each of them are based on an approximate description of the 2D inversion layer (neglecting degeneracy, quantization or accounting for only one single subband), leading to an inaccurate estimation of the injection velocity. Conventional numerical Poisson-Schrodinger simulations are not suitable with compact models for device simulations and performance predictions.

In this chapter, we explain how strain alters the valence and conduction band of SiGe, SiGeC and strained-Si. Electronic properties and carrier transport are also the topics of this chapter. After a discussion of the band structure of Si/SiGe heterostructures, physical mechanisms of electron mobility enhancement in uniaxial stressed MOSFETs and impact of uniaxial stress engineering in ballistic regime are discussed. Key differences for process-induced uniaxial vs. substrate-induced biaxial stressed Si and Ge channel MOSFETs are discussed. An attempt will also be made to confirm the differences for biaxial and uniaxial stress, and explain the physical mechanism responsible.

## 4.1 Substrate-induced Strained-Si

Uniaxial stress has become the preferred method to strain the Si lattice and has been implemented into several 90-nm logic technologies [3, 11, 16]. However, during the last 15 years, major focus has been placed on tensilely strained biaxial substrate-strain induced strained-Si [17]. The biaxial tension (substrate strain) in strained-Si grown on relaxed-SiGe substrates splits the degeneracy in both the conduction and valence bands and reduces the bandgap

[18]. For n-MOSFETs, the tensile strain in pseudomorphic Si on relaxed-SiGe splits the six-fold degeneracy of the conduction band minimum, resulting in an increased electron mobility due to the lower in-plane effective mass and reduced intervalley scattering.

Band alignments in heterostructures giving rise to band offsets or band discontinuities, which are very important for junction device applications. For bipolar heterostructure devices, the band offset determines the forward injection efficiency. On the other hand, it controls the fraction of charge carriers confined in the high-mobility channel in MOSFETs. For an atomically sharp interface, the band discontinuities are sharp on an atomic length scale in contrast to the much larger length in band bending associated with depletion layers. Therefore, the band discontinuities are more fundamental in nature and are used as boundary conditions in the solution of Poisson's equation to yield the band bending.

It is expected that the tensile strain, which is perpendicular to the (110) surface, induces the energy splitting between the four-fold and the two-fold valleys, and the energy level four-fold valley becomes lower than that of the two-fold valley. As a result, the electron mobility along the direction can be enhanced because of lower conductivity mass of the four-fold valley along the direction and the reduced intervalley scattering, which is similar to electrons in the (100) strained-Si channel [19]. This increased mobility has been demonstrated in n-MOSFETs by various workers [20, 21, 22] and has been shown to provide significant drive current enhancement in deep submicron n-MOSFETs in both bulk and SOI technologies [8, 23]. This is attributed to the increased source-side carrier velocity that is enabled by the larger low-field mobility [24, 25] in strained-Si compared to bulk-Si.

The electron mobility is enhanced in a strained-Si layer grown on top of a thick, relaxed-SiGe due to the lattice mismatch (which can be controlled by the Ge mole fraction), the thin silicon layer appears to be "stretched" (under biaxial tension) [26]. The origin of the electron mobility enhancement can be explained by considering the six-fold degeneracy in the conduction band. It is known that the perpendicular effective mass is much lower than the longitudinal one. Therefore, this carrier redistribution and reduced intervalley scattering enhance the electron mobility. The hole depends on the strain mainly due to redistribution of holes between light and heavy valleys, and changes the effective masses in these valleys.

In the deformation potential theory, the strains are considered to be relatively small. The change in energy of each carrier subvalley, caused by the deformation of the lattice, is a linear function of the strain. For silicon, three subvalleys for electrons (which are applied to three two-fold subvalleys in the conduction band) and two subvalleys for holes (which are applied to heavy-hole and light-hole subvalleys in the valence band) are considered.

To quantify the mobility enhancement, changes in the conductivity effective mass and scattering are taken into account, both of which depend on the strain-altered valence band. Biaxial tension in strained-Si grown on relaxed-

$Si_{1-x}Ge_x$ virtual substrates alters the Si band structure by splitting the subband degeneracy in both the conduction and valence bands and reducing the bandgap [27]. To explain the field dependence of the mobility in biaxial tensile strain, Fischetti et al. [28] showed the reduction in hole mobility enhancement with vertical field results from the quantization associated with the confining surface potential of the MOSFET. Significant scattering reduction requires $>$ 1 GPa of stress to split the band 60 meV, which is consistent with the biaxial substrate hole mobility experimental data. For biaxial stress since there is no mass improvement, the mobility enhancement only results from reduced scattering and thus requires $\sim$ 25-30% Ge ($>$ 1 GPa stress) [29].

### 4.1.1 Energy Gap and Band Structure

The band structure determines the electronic and optical properties of semiconductors. Transport properties, the bandgap, and band shapes control optical absorption and photoluminescence. The knowledge of band alignment is essential to determine the band offsets, which in turn play an important role in the design of heterostructure devices. The high lattice mismatch in SiGe heterostructures significantly modifies the band structures giving rise to many complexities in band structure calculations. The effective mass approximation model, valid for lattice matched heterostructures such as GaAs-AlGaAs, does not hold good for a strained layer system, which requires a microscopic formalism taking into account the precise atomic positions and potential variations in the crystal. In the empirical pseudopotential (EPP) approach used for bulk semiconductors, [30], the empirically known band structure properties are reproduced by choosing a suitable pseudopotential. A perturbation potential is then constructed for each atom in the unit cell to account for the lattice strain introduced into the bulk crystal. The consideration of microscopic variation of potential due to individual atoms gives rise to a clear description of bulk momentum mixing resulting from the scattering of electrons from the misplaced atoms. The scheme accurately describes the states at and around the conduction and valence band edges. Therefore, the pseudopotential method has been found to be successful in predicting various optical and transport properties of strained-SiGe layers. The approach has been found to have limitations in the presence of impurity atoms when the deviations from the bulk-like structure are too great due to structural relaxation.

On the other hand, the local density approximation (LDA) [31, 32] and abinitio pseudopotential calculations provide a microscopic description of SiGe alloys without any reference to the bulk crystal properties and describe the properties of a particular atomic species regardless of its local environment. The LDA schemes, which include the microscopic potential variation in the interface region, give a very accurate description of the valence band through minimization of ground-state energy and predict structural relaxation very precisely at the cost of an increased computational effort. However, the value of the bandgap is vastly underestimated due to an inadequate description of

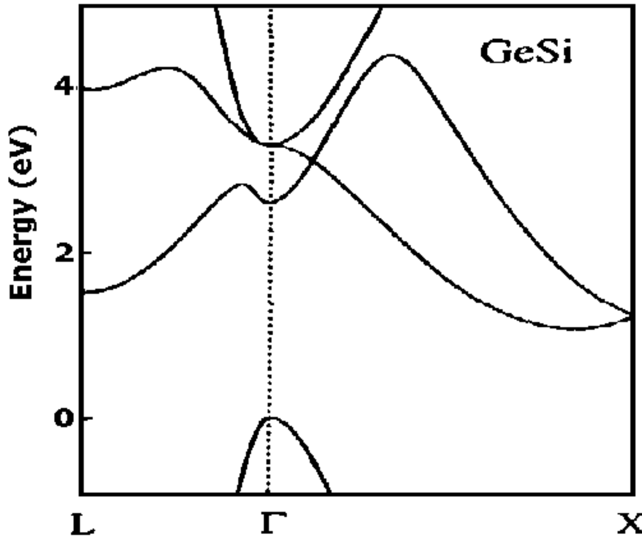



**FIGURE 4.1**
The valence band structure of Si and Ge showing lower hole effective masses and very large spin-orbit splitting in Ge as compared to Si.

the excited states of the system in the LDA approach. Both the EPP and LDA methods have contributed significantly to an understanding of the band structure and transport properties of strained-SiGe layer heterostructures. The suitability of a particular method depends upon the particular structure, the nature of the system and specific properties that need to be studied. In the following, we present the band structure and transport properties of group-IV alloy semiconductor heterostructures and strained layer superlattices.

An important consideration for any device utilizing the silicon heterostructures is the internal band alignment of the structure due to the bond interface. Previous reports of similar strained-Si structures grown on relaxed-SiGe buffers with similar levels of strain have shown that the strained-Si channel behaves as an electron well with a depth of approximately 130-200 meV [33]. The bandgap of such strained-Si layers has also been determined to be approximately 1 eV. Based on these previous determinations of the band alignment of strained-Si relative to bulk-Si, one can speculate on the likely internal band alignment of the SSOI heterostructures.

The empirical pseudopotential formalism has been used by treating the SiGe alloy as a "virtual crystal" having the interpolated lattice constant, potential, and so on. The band structure of unstrained-$Si_{0.5}Ge_{0.5}$ alloy using the VCA is shown in Fig. 4.1 along the symmetry lines L-Γ-X. The bandgap of unstrained bulk SiGe alloy was measured at 4.2 K [34]. The excitonic bandgap varies

smoothly with Ge content from the Si gap at 1.155 eV to the Ge gap at 0.740 eV. For the $Si_{1-x}Ge_x$ alloy, the conduction band structure is Si-like up to 85% Ge with minima at the X-point. However, it exhibits Ge-like character with minima at the L-point for Ge fractions above 85%. The experimentally measured values could be fitted well by the quadratic equation [35]

$$E_{\mathrm{g}}(x) = (1.155 - 0.43x + 0.0206x^2) \;\; \text{eV for } 0 < \text{x} < 0.85 \tag{4.1}$$

and

$$E_{\mathrm{g}}(x) = (2.010 - 1.27x) \;\; \text{eV for } 0.85 < \text{x} < 1 \tag{4.2}$$

There is a large body of experimental data on the effects of compressive strain on the bandgap, but comparatively less for the effects of tensile strain, i.e., for $Si_{1-x}Ge_x$ grown on a Ge substrate or Si grown on a relaxed-$Si_{1-x}Ge_x$ substrate. However, theoretical calculations of the bandgap of arbitrary strained alloy layers have been reported by several authors [36, 37, 38].

For compressively strained-$Si_{1-x}Ge_x$ on Si substrate, the theoretical results of People [36, 37] using deformation potential theory, indicated by the hatched lines in Fig. 4.2, are in good agreement with the experimental data. The bandgap of compressively strained-$Si_{1-x}Ge_x$ layer can be approximated by the relationship [39]

$$E_{\mathrm{g}}(x, T) = E_{\mathrm{o}}(T) - 0.96x + 0.43x^2 - 0.13x^3 \tag{4.3}$$

where $E_{\mathrm{g}}$ is the difference between the lowest lying conduction band and the highest lying valence band edge for a $Si_{1-x}Ge_x$ strained layer grown on Si and $E_{\mathrm{o}}(T)$ is the bandgap of bulk-Si.

As seen from Fig. 4.2, the theoretically reported values for the bandgap of strained-SiGe alloys on a Ge or a $Si_{0.5}Ge_{0.5}$ substrate differ quite substantially. The solid line was reported by People [36, 37] by considering the strain induced shift and splitting of conduction and valence bands using the deformation potential parameters of bulk-Si for x < 0.85 and those for bulk Ge for x > 0.85. However, the use of deformation potential parameters by linearly extrapolating the respective values of Si and Ge yielded a different result, as shown by the dashed-dotted lines in Fig. 4.2. On the other hand, the calculated values of bandgap based on local density fluctuations and ab-initio pseudopotentials [38] are shown by the dashed line. Therefore, experimental values of bandgap are necessary for tensile strained-SiGe alloys grown on higher lattice constant substrates.

Lattice strain alters the band structure of a semiconductor by either shifting it in energy, distorting it, removing degeneracy effects, or any combination of the three. Hydrostatic strain shifts the energy level of a band. On the other hand, a uniaxial or biaxial strain splits the degenerate bands. When a thin layer of silicon is grown on top of a thicker layer of silicon germanium, the lattice mismatch pulls the silicon atoms apart. The conduction band in silicon contains six equal ellipsoids, and the strain causes four of them to go up in

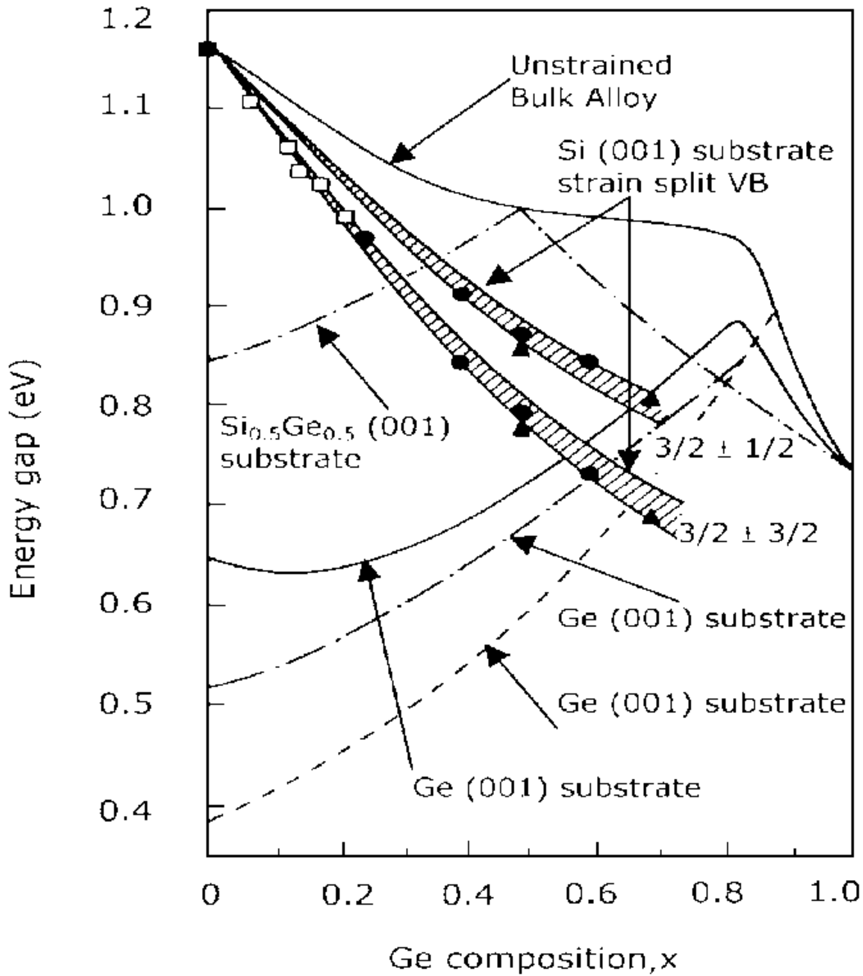



**FIGURE 4.2**
Fundamental indirect bandgap of strained-$Si_{1-x}Ge_x$ alloys in comparison with bulk alloy. The full (open) symbols are the results of photocurrent measurements at T = 90 K for pseudomorphic $Si_{1-x}Ge_x$ films on Si(001) substrates. The hatched areas between the full lines show the theoretical results for transitions involving heavy-holes ($\pm3/2$) and light-holes ($\pm1/2$). Also shown are the calculated data for pseudomorphic films on either $Si_{0.5}Ge_{0.5}$ or Ge substrates.

energy and two to go down in energy. Uniaxial strain splits the band into two-fold degenerate bands at lower energy plus a nondegenerate band at higher energy while keeping the average energy over the three bands equal. The energy shift due to strain does not alter the shape of the conduction band and keeps the effective mass unaltered. However, for multivalleyed semiconductors like Si, Ge and $Si_{1-x}Ge_x$ alloys the carrier population in individual bands may be altered. On the other hand, the strain has a significant effect on the shape of the valence band structure, resulting in a change in effective mass and mobility. So the piezoresistance effect is much more predominant in p-type semiconductors than in n-type semiconductors. Additionally, reduced symmetry of the band structure leads to a shift in energy of the valence band.

In case of an unstrained system, i.e., for a lattice-matched heterojunction, the band offsets can be simply derived from the band lineup at the interface. The calculation is more complicated in the case of strained heterostructures with group-IV alloys. Besides the normal band offsets, the splitting of bands due to uniaxial strain and additional shifts due to hydrostatic strain need to be computed. It should be noted that the band discontinuities in strained heterostructures are well defined only if the interface is pseudomorphic so that the in-plane lattice constant is continuous across the interface. The best theoretical study on the estimation of heterojunction discontinuities has been carried out by van de Walle and Martin [38].

As discussed above, the strain alters the position of the atoms in a crystal, and in a tight-binding (TB) model, both the overlap energies between orbitals sitting on neighboring atoms and the on-site orbital energies are altered because of this. Using TB model, Rahman et al. [40] have calculated the strain in strained-Si p-MOSFETs where both the uniaxial compressive and biaxial tensile strains are involved. Fig. 4.3 shows the two different strained p-MOSFET geometries compared. In both cases, relaxed-SiGe induces strain in the pseudomorphic Si channel. Their lattice mismatch causes biaxial tensile strain in Si channel along the SiGe-Si interfacial plane and uniaxial compressive strain normal to it. The holes are quantum mechanically confined along thickness direction, z. It may be noted that in these two geometries, the relative orientation of the uniaxial strain axis and confinement axis are different.

This relative orientation of strain with quantum confinement has a profound effect on the transport properties at high gate bias, as has already been observed experimentally [11]. The combined effect of strain and quantum confinement on the Si valence subbands is presented in Fig. 4.4. In bulk-Si, quantum confinement is absent and biaxial tensile strain removes the degeneracy of heavy hole (HH)-light hole (LH) band at Γ. As a result, the LH band with better transport properties rises in energy and is primarily populated by holes. In Fig. 4.4 we observe that this strain advantage is nullified by quantum effects in strained-Si on SiGe devices (left), and the top subband again becomes HH-like with poor transport properties. However, for the strained-Si channel in SiGe S/D devices (right), the top subband remains LH-like (large curvature) even at strong quantum confinement. This bandstructure advan-

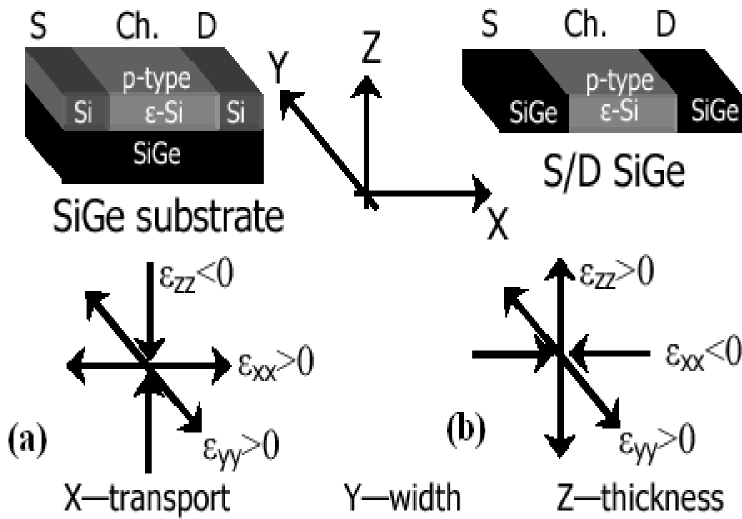



**FIGURE 4.3**
Two strained-Si p-MOSFET geometries are compared. (a) SiGe substrate (substrate-induced strain) and (b) SiGe source/drain (process-induced stress) devices. Lattice mismatch between SiGe and Si causes biaxial tensile strain ($\epsilon > 0$) in Si, parallel to the SiGe-Si interface, and compressive strain ($\epsilon < 0$) normal to the interface. For the SiGe substrate device, compressive strain is along the quantum confinement direction of inversion layer holes, while in the SiGe S/D device it is along the transport. After A. Rahman et al., *IEEE IEDM Tech. Dig.*, 2004(139-142). With permission.

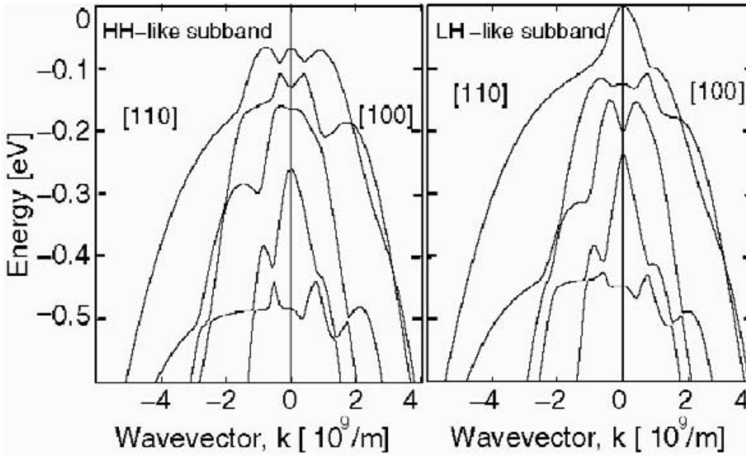



**FIGURE 4.4**
Tight-binding simulations of strain effects in quantized valence subbands of Si. Thickness of Si-(001) channel is 24 atomic layers ($\sim$ 3.5 nm). Strain inducing SiGe regions contain 25% Ge. Left: In SiGe substrate device, top hole subband in Si is HH type (small curvature). Right: Top hole subbands of Si in SiGe S/D device is LH type (large curvature). In presence of strong quantum confinement, a clear bandstructure advantage of SiGe S/D device is visible. After A. Rahman et al., *IEEE IEDM Tech. Dig.*, 2004(139-142). With permission.

tage has led to 25% more $I_{on}$ in SiGe S/D device compared to its counterpart.

The conduction band structure of $Si_{1-x}Ge_x$ is silicon-like below x = 0.85. The conduction band minimum in Si occurs at the $\Delta$-point between the $\Gamma$- and X-valleys and is strongly affected by a tetragonal strain. The L conduction band minimum of Ge has a four-fold spatial degeneracy, which is not lifted by a tetragonal strain. Since the L-valley for Ge becomes important only for very high Ge concentration, the six-fold degenerate bands at the $\Gamma$-point are considered here for $Si_{1-x}Ge_x$ alloys of high practical interest. The six ellipsoidal energy shells located at equivalent positions in *k*-space in <001> directions are shown in Fig. 4.5(a). The strain has no apparent effect on the ellipsoidal shape of the energy surface but causes a shift in energy [41, 42]. So unlike the valence band cases, the effective mass of each band remains unchanged [38]. The relative shift in energy in the case of a compressively strained film is shown in Fig. 4.5(b) where [001] conduction bands (two-fold degenerate $\Delta_2$ bands) move up in energy, while the [010] and [100] bands (four-fold degenerate $\Delta_4$ bands) are shifted downwards in energy. Fig. 4.5 shows the constant energy surface diagram indicating the size change of each valley to a smaller one for an upward energy shift and to a larger one for a downward

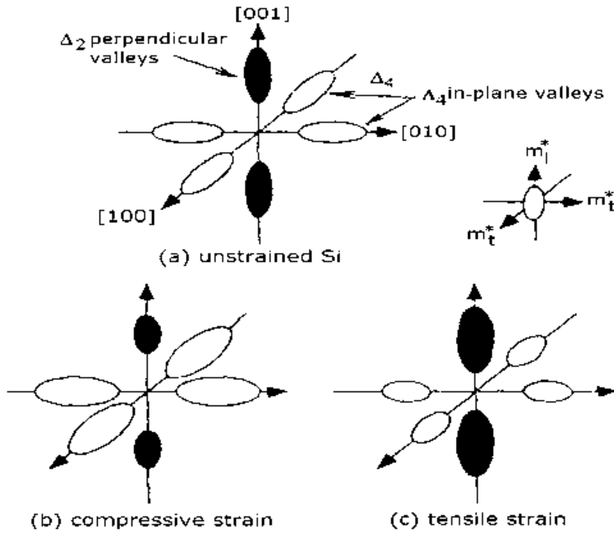



**FIGURE 4.5**
Constant energy surfaces for the Si conduction band. The size change of each valley in the constant energy surface diagram indicates a shift up (smaller) or down (larger) in energy. After T. Manku and A. Nathan, *IEEE Trans. Electron. Dev.*, Vol. 39, 1992(2082-2089). With permission.

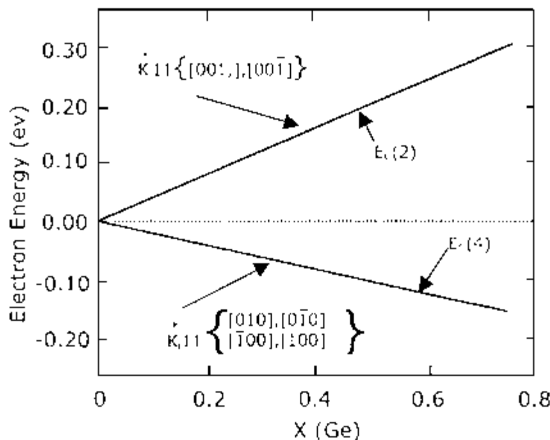



**FIGURE 4.6**
$\Delta_1$ conduction band splitting (uniaxial components only) for bulk-$Si_{1-x}Ge_x$ alloys grown on Si(001) substrates. After R. People, *Phys. Rev. B*, Vol. 32, 1985(1405-1408). With permission.

shift. In the case of tensile strained film (film grown pseudomorphically on a substrate with larger lattice constant), the direction of motion of the split levels is reversed. Here, the $\Delta_2$ band moves lower in energy relative to the $\Delta_4$ band as shown in Fig. 4.5(c).

The splitting of the conduction band minima for compressively strained $Si_{1-x}Ge_x$ films as a function of Ge concentration is shown in Fig. 4.6. The energy difference between the separated bands due to splitting depends on the magnitude of the strain and can be calculated by the deformation potential theory [43, 44]. Deformation potential theory has also been used by van de Walle and Martin [38] to give expressions for the perturbed conduction band levels of Si and Ge.

In order to determine the band alignment for a given $Si_{1-x}Ge_x$/Si heterointerface, the bandgap of the constituent layers as well as the band discontinuity, $\Delta E_c$ or $\Delta E_v$, must be known. Strain-induced modification of Si/SiGe films is found to have a significant impact on the band structure and carrier transport. When a thin film with a larger lattice constant (e.g., $Si_{1-x}Ge_x$) is grown on a smaller lattice constant substrate (e.g., Si), the film maintains the in-plane lattice constant of the substrate and is under a biaxially compressive strain. This is known as the type-I band alignment where almost the entire band offset occurs in the valence band while the band offset in the conduction band is very small. This type of structure is favorable for hole confinement and has been exploited in novel heterostructure devices, namely p-MOSFETs, p-MODFETs, and HBTs [45].

Similarly, a smaller lattice constant Si epilayer will be under biaxial tension when grown on a larger lattice constant relaxed-$Si_{1-x}Ge_x$ substrate. In this case, type-II band offset occurs and the structure has several advantages over the more common type-I band alignment, as a large band offset is obtained in both the conduction and valence bands, relative to the relaxed-$Si_{1-x}Ge_x$ layer [43]. This allows both electron and hole confinements making it useful for both n- and p-type devices for strained-Si/SiGe-based CMOS technology. Since strained-Si provides larger conduction and valence band offsets and does not suffer from alloy scattering (hence mobility degradation) [46], a significant improvement in carrier mobility can be achieved. Strained-Si is more difficult to grow than strained-$Si_{1-x}Ge_x$, since bulk-$Si_{1-x}Ge_x$ substrate is not available and growth of relaxed-$Si_{1-x}Ge_x$ without forming a large concentration of defects due to dislocations was difficult. However, the ability to achieve both nMOS and pMOS devices using strained-Si provides a promising alternative for the next-generation high-performance SiGe-CMOS technology (see, for example, reviews by Schaffler [47] and Maiti et al. [18] and references therein).

Bandgap narrowing due to a heavy doping effect changes the physical constants used in the minority carrier transport equation and is crucial for accurate determination of device properties. Jain and Roulston [48] reported general closed-form equations for bandgap narrowing for n- and p-type Si, Ge, GaAs, and $Si_{1-x}Ge_x$ strained layers. The equations are derived by identifying the exchange energy shift of the majority band edge, correlation energy shift

of the minority band edge, and impurity interaction shifts of the two band edges. The bandgap narrowing (for N > $10^{18}$ cm$^{-3}$) is given by

$$\Delta E_{\mathrm{g}} = A_1\left(\frac{N}{10^{18}}\right)^{1/3} + A_2\left(\frac{N}{10^{18}}\right)^{1/4} + A_3\left(\frac{N}{10^{18}}\right)^{1/2} \tag{4.4}$$

where $A_1$, $A_2$ and $A_3$ for n- and p-type Si, Ge and GaAs are given in reference [49]. For p-type $Si_{1-x}Ge_x$ alloy with a Ge content of less than 0.3, the bandgap narrowing is given by

$$\begin{aligned}\Delta E_{\mathrm{g}} = 11.07(1-0.35x)\left(\tfrac{N}{10^{18}}\right)^{1/3} + 15.17(1-0.54x)\left(\tfrac{N}{10^{18}}\right)^{1/4} + \\ 5.07(1+0.18x)\left(\tfrac{N}{10^{18}}\right)^{1/2}\end{aligned} \tag{4.5}$$

Assuming that the linear dependence on Ge concentration is independent of doping, the effects of bandgap reduction due to Ge concentration and bandgap narrowing due to heavy doping are separated [50]:

$$\Delta E_{\mathrm{g,eff}} = \Delta E_{\mathrm{g,dop}} + \Delta E_{\mathrm{g,Ge}} \tag{4.6}$$

Assuming a $\Delta E_{\mathrm{g,dop}}$ of the form

$$\Delta E_{\mathrm{g,dop}} = A + B\log\left(\frac{N_{\mathrm{A}}}{10^{18}}\right) \tag{4.7}$$

and a linear dependence of $\Delta E_{\mathrm{g,Ge}}$ on $x$, a three-parameter best fit was found to be

$$\Delta E_{\mathrm{g,eff}} = 28.6 + 27.4\log_{10}\left(\frac{N_{\mathrm{A}}}{10^{18}}\right) + 688x \tag{4.8}$$

where $N_A$ is the base doping and $x$ is the Ge mole fraction and Eqn. 4.8 is valid only for doping levels beyond $10^{18}$ cm$^{-3}$. The first two terms represent bandgap narrowing due to heavy doping and the last term is the bandgap reduction due to the Ge contribution. $\Delta E_{\mathrm{g,eff}}$ is not a measure of the actual bandgap reduction but the effective (apparent) bandgap reduction relevant for minority carrier concentration and thus electron transport across the $Si_{1-x}Ge_x$. The apparent bandgap is larger than the true bandgap due to valence band filling in the degenerately doped semiconductor.

In devices, impurity scattering plays an important role in limiting the achievable mobility. The Brooks-Herring and Cornwell-Weisskopf approaches [51] are used in modeling three-dimensional charged impurity scattering including the effect of screening by carriers. The drift mobility and velocity field characteristics of a bulk semiconductor are limited by intrinsic lattice scattering through interaction between the phonons and the carriers, impurity scattering in doped semiconductors and alloy scattering in semiconductor

alloys. Phonon scattering is contributed by acoustic phonons and optical phonons, each with its characteristic dispersion relations. These are specified by the relevant deformation potentials, which determine the rate of scattering from an initial momentum to a final momentum, with the states in the same or a different valley, i.e., intravalley scattering or intervalley scattering. The actual scattering rate depends on the band structure details, the number of equivalent and nonequivalent valleys and their energy separation. The interaction potential for impurity scattering is related directly to the screened Coulombic potential while for alloy scattering an alloy potential whose effect depends on the mole fractions of the mixture is used. The actual mobility measured in a device is smaller than the bulk mobility, however, because of the presence of interfaces invariably present with their defects and charges.

Computations of band properties in relation to determination of scattering processes and bulk mobility of carriers have been made by Hinckley and Singh [52]. Acoustic phonon and optical phonon scattering rates and their dependence on energy at room temperature have also been computed. At very low energies, the optical phonon emission process is absent and the acoustic phonon scattering rate ($W_{\rm ac}$) and nonpolar optical phonon absorption rate (NPOA) are comparable. It can be concluded that (a) the nonpolar optical (NPO) scattering rate ($W_{\rm np}$) is greater than the acoustic phonon scattering rate, (b) at low energies the impurity scattering rate may exceed $W_{\rm np}$, and (c) $W_{\rm alloy}$ for an alloy potential of 0.8 eV exceeds $W_{\rm np}$ and phonon scattering rates for Ge are smaller typically by a factor of about two. This combined with a smaller effective mass for holes in Ge gives higher mobility.

For MOSFET applications, the mobility is mainly dominated by the phonon-phonon scattering, however, for low temperature transport study, instead of phonon-phonon scattering, the mobility is determined by the Si/SiGe interface roughness scattering, impurities scattering, alloy scattering, etc. Fig. 4.7 shows the schematic of the universal mobility curve of Si MOSFETs at room temperature. Carrier mobility is limited by Coulomb, phonon, and surface roughness scattering. Quantifying the improvement due to scattering has been difficult using acceptable scattering parameters, but reduced scattering is still believed to account for the mobility enhancement [53].

The general formalism for computing momentum relaxation time in the case of acoustic phonon scattering $\tau_{\rm ac}$ and zero-order optical phonon scattering $\tau_{\rm op}$ is given by [51]

$$\tau_{\rm ac} = \frac{\pi \hbar^4 c_{\rm L}}{2^{1/2} \Xi^2 m^{*3/2} k_{\rm B} T E_k^{1/2}} \tag{4.9}$$

and

$$\tau_{\rm op} = \frac{2^{1/2} \pi \hbar^3 \omega_{\rm o} \rho}{D_{\rm o}^2 m^{*3/2}} \left\{ n(\omega_{\rm o})(E_k + \hbar\omega_{\rm o})^{1/2} + [n(\omega_{\rm o}) + 1](E_k - \hbar\omega_{\rm o})^{1/2} \right\}^{-1} \tag{4.10}$$

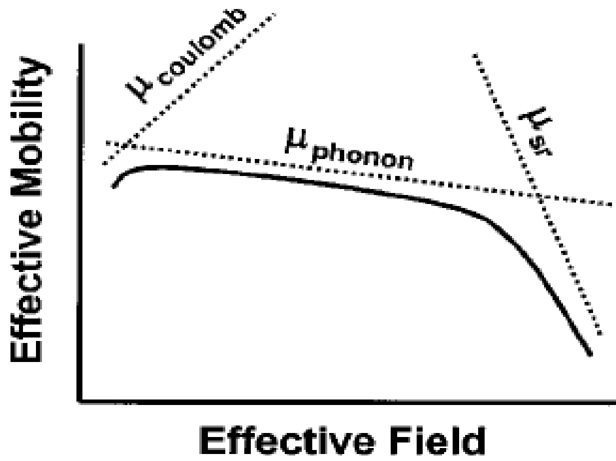



**FIGURE 4.7**
Schematic of the universal mobility curve of Si MOSFETs at room temperature. Carrier mobility is limited by Coulomb, phonon, and surface roughness scattering.

where $\Xi$ is the deformation potential, $D_o$ is the interaction constant, $n(\omega_o)$ is the Bose-Einstein occupancy function with optical phonon frequency $\omega_o$, $E_k$ is the carrier energy, $\rho$ is the mass density, $\hbar$ is the reduced Planck's constant and $c_L$ is the elastic constant of the lattice in the longitudinal mode. The expressions show that the acoustic phonon limited mobility increases as $T^{-1}$. The variation of optical phonon limited mobility arising from the Bose-Einstein occupancy function $n(\omega_o)$ shows a faster increase with inverse temperature.

The expression for momentum relaxation time of charged impurity scattering in the Brooks-Herring model is given by [51]

$$\tau_{\text{imp}} = \frac{2^{9/2}\pi\epsilon^2 m^{*1/2} E_k^{3/2}}{Z^2 e^4 N_\text{I}} \left[ \log\left(1 + \frac{8m^* E_k}{\hbar^2 q_\text{o}^2}\right) - \frac{1}{1 + \hbar^2 q_\text{o}^2 / 8m^* E_k} \right]^{-1} \quad (4.11)$$

where $q_o$ is the reciprocal screening length, $N_I$ is the number of impurity centers per unit volume and $Z$ is the atomic number. Mobility limited by impurity scattering decreases with temperature, varies inversely with doping concentration, and increases with carrier concentration. The scattering rate depends on the separation between the carrier supply layer and the channel and is therefore high for accumulation mode devices when the dopants are separated from the channel. This effect is exploited in modulation-doped structures and may be computed in a manner similar to that employed for finding the effect of oxide charge.

The band structure of an alloy is described using the virtual crystal approximation assuming a uniform distribution of the constituents. Fluctuations from uniformity give rise to local changes in the potential experienced by the carriers, and hence scattering. The scattering rate is proportional to the magnitude of the scattering potential $V_{\mathrm{a}}$ and the product $x$(1-$x$) where $x$ is the Ge mole fraction in the binary alloy. The alloy scattering limited mobility in a $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ alloy is given by [54]

$$\mu_{\mathrm{alloy}} = \frac{e\hbar^3}{m_{\mathrm{t}}^{*2}\Omega_{\mathrm{o}}V_{\mathrm{a}}^2 x(1-x)}\left(\frac{16}{3b}\right) \tag{4.12}$$

where $m_{\mathrm{t}}^*$ is the transverse effective mass, $\Omega_{\mathrm{o}}$ is the atomic volume, $V_{\mathrm{a}}$ is the alloy disorder potential and $b$ is the variational parameter given by

$$\mathrm{b} = \left(\frac{33m_{\mathrm{l}}^* e^2 n}{8\epsilon\hbar^2}\right)^{1/3} \tag{4.13}$$

where $m_{\mathrm{l}}^*$ is the longitudinal effective mass and $n$ is the carrier concentration. Thus, the alloy scattering mobility has dependence on the masses and the carrier concentration. It is presumed that alloy scattering limited mobility in relaxed- and strained-SiGe materials would be similar. An estimation of the alloy disorder potential $V_{\mathrm{a}}$ is important as it is squared in the formula.

### 4.1.2 Electron Mobility

Strained-Si has been studied for nearly 15 years and the physics behind the inversion layer mobility enhancement has been reported [19, 27, 46, 55]. Mobility improvement in strained-Si takes place mainly due to:

(a) the reduction in the carrier conductivity effective mass; and

(b) the reduction in the intervalley phonon scattering rates [56, 57, 58].

Both facts are a consequence of the splitting of the six-fold degeneracy in the Si conduction band minimum caused by the strain [59]. However, the effect on the mobility of electrons, the main carrier in n-channel devices, is different than that of hole mobility, the main carrier in p-channel devices. While it is relatively easy to get close to a factor of 2 increase in electron mobility, it is more complicated to get a similar increase in hole mobility. Though there has been much focus on reduced in-plane mass to increase mobility, increasing the out-of-plane mass for electrons and holes is now understood to be equally important for maintaining the mobility enhancement at high vertical fields [60].

For electron transport in bulk-Si at room temperature, the conduction band is comprised of six degenerate valleys. Strain removes the degeneracy between the four in-plane valleys ($\Delta_4$) and the two out-of-plane valleys ($\Delta_2$) by splitting them in energy. The conduction band structure of $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ is silicon-like below $x = 0.85$. The conduction band minimum in Si occurs at the $\Delta$-point between the $\Gamma$- and X-valleys and is strongly affected by a tetragonal strain.

The L conduction band minimum of Ge has a four-fold spatial degeneracy, which is not lifted by a tetragonal strain. Since the L-valley for Ge becomes important only for very high Ge concentration, the six-fold degenerate bands at the Γ-point are considered here for $Si_{1-x}Ge_x$ alloys of high practical interest. The six ellipsoidal energy shells located at equivalent positions in $k$-space in <001> directions are shown in Fig. 4.5(a). The strain has no apparent effect on the ellipsoidal shape of the energy surface but causes a shift in energy [41, 42]. So unlike the valence band cases, the effective mass of each band remains unchanged [38]. The relative shift in energy in the case of a compressively strained film is shown in Fig. 4.5(b) where [001] conduction bands (two-fold degenerate $\Delta_2$ bands) move up in energy, while the [010] and [100] bands (four-fold degenerate $\Delta_4$ bands) are shifted downwards in energy.

In the case of tensile strained film (film grown pseudomorphically on a substrate with larger lattice constant), the direction of motion of the split levels is reversed. Here, the $\Delta_2$ band moves lower in energy relative to the $\Delta_4$ band as shown in Fig. 4.5(c). The lower energy of the $\Delta_2$ valleys means that they are preferentially occupied by electrons. The electron mobility partly improves via a reduced in-plane and increased out-of-plane $m^*$ due to the favorable mass of the $\Delta_2$ valleys, which results in more electrons with an in-plane transverse effective mass ($m_t = 0.19\ m_o$) and out-of-plane longitudinal mass ($m_l = 0.98\ m_o$).

The carrier mobility is given by

$$\mu = \frac{q\tau}{m^*} \tag{4.14}$$

where $1/\tau$ is the scattering rate and $m^*$ is the conductivity effective mass. Strain enhances the mobility by reducing the conductivity effective mass and/or the scattering rate. For unstressed bulk-Si, the total electron conductivity mass, $m^*$, is obtained by adding the contributions of the six degenerate valleys and is given by

$$m^* = \left[\frac{1}{6}\left(\frac{2}{m_l}\right) + \left(\frac{4}{m_t}\right)\right]^{-1} \tag{4.15}$$

The effective mass for any direction is the reciprocal of the curvature of the electron energy function in that direction. Consequently, the effective mass of each ellipsoid is anisotropic, with the transverse mass (perpendicular to the axis) given by $m_t = 0.19\ m_o$ being significantly smaller than the longitudinal mass (parallel to the axis) given by $m_l = 0.98\ m_o$, where $m_o$ is the free electron mass. For electrons, both mass and scattering changes are generally accepted as important for mobility enhancement [53]. The electron scattering also reduces due to the conduction valleys splitting into two sets of energy levels, which lowers the rate of intervalley phonon scattering between the $\Delta_2$ and $\Delta_4$ valleys. However, for holes, only mass change due to band warping and repopulation plays a significant role, as the strain-induced valence band splitting is smaller than that for the conduction band [55].

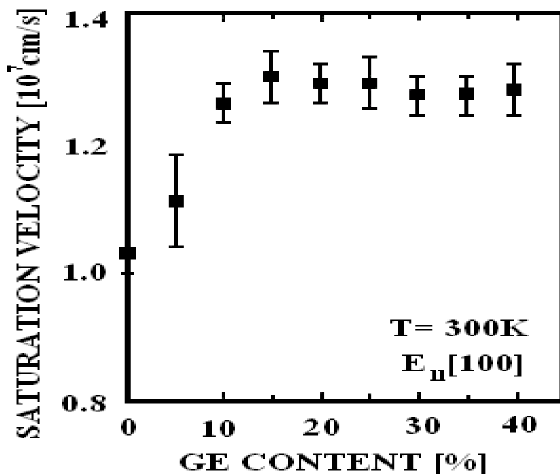



**FIGURE 4.8**
In-plane drift velocity in strained-Si on SiGe for an electric field of $10^5$ V/cm vs. Ge content of the substrate (error bars: standard deviation of different Monte Carlo calculations). After Th. Vogelsang and K. R. Hofmann, *Proc. IEEE DRC*, 1992(34-35). With permission.

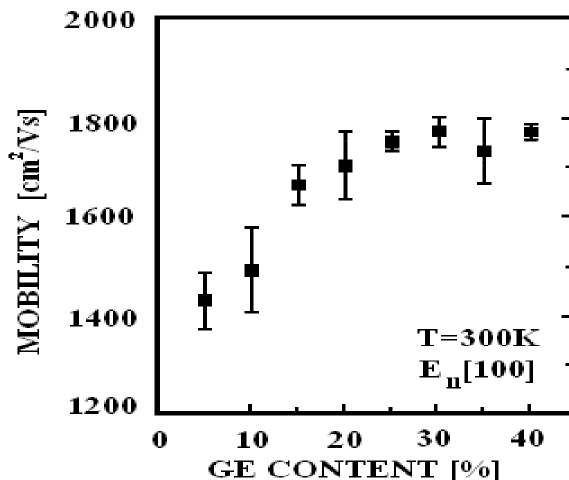



**FIGURE 4.9**
Low field in-plane drift mobility of strained-Si on SiGe at E = 1 kV/cm vs. Ge content of the substrate (error bars: standard deviation of different Monte Carlo calculations). After Th. Vogelsang and K. R. Hofmann, *Proc. IEEE DRC*, 1992(34-35). With permission.

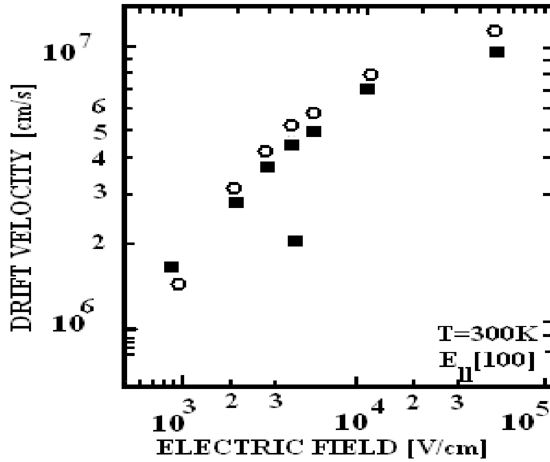



**FIGURE 4.10**
In-plane velocity-field characteristics of strained-Si on SiGe. Open circles: substrate with 25% Ge content. Squares: unstrained-Si. After Th. Vogelsang and K. R. Hofmann, *Proc. IEEE DRC*, 1992(34-35). With permission.

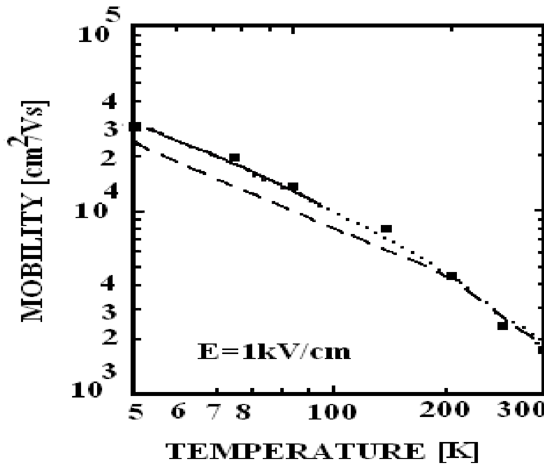



**FIGURE 4.11**
Calculated temperature dependence of in-plane drift mobility of strained-Si/$Si_{0.75}Ge_{0.25}$ (squares) compared to experimental Hall mobility in Si/$Si_{0.75}Ge_{0.25}$ modulation doped structure for two different samples of [1] dashed and dotted lines and [2] solid line. 1. F. Schaffler et al., *Semicond. Sci. Technol.*, 7, 260 (1992). 2. Y. J. Mii et al., *Appl. Phys. Lett.*, 59, 1611 (1991). After Th. Vogelsang and K. R. Hofmann, *Proc. IEEE DRC*, 1992(34-35). With permission.

Vogelsang and Hofmann [61] have investigated the electron drift velocities, mobilities, mean carrier energies and valley populations for in-plane electric fields from $1\times10^3$ to $1\times10^5$ V/cm, for temperatures between 300 K and 50 K, and Ge contents of the substrate up to 40% (see Figs. 4.8 and 4.9). A substantial increase of both low-field mobility and of high-field drift velocity was observed when the Ge content, x, of the substrate is raised from 0 to 40%. Both quantities saturate at a comparatively small x values. At room temperature, the low-field mobility increases from 1420 cm$^2$/V-s in unstrained Si to a saturated value of about 2000 cm$^2$/V-s at a Ge content of 30%. The high-field drift velocities show considerable improvement relative to unstrained-Si, e.g., 16% at an intermediate field of $1\times10^4$ V/cm for 25% Ge (see Fig. 4.10), in the whole field range up to about $1\times10^5$ V/cm, where they saturate just above the unstrained-Si saturation velocity $1\times10^7$ cm/s.

The calculated low-field (1 kV/cm) drift mobility data agree very well (over the whole calculated temperature range) with the highest Hall mobilities measured by [62, 63] on thin (15-20 nm) modulation-doped Si channels on relaxed-$Si_{1-x}Ge_x$ (x = 25-30%) buffers. These results demonstrate significant improvements of the in-plane electron drift velocity in strained-Si on $Si_{1-x}Ge_x$ compared to bulk-Si in the low-field and in the high-field region both at 300 K and 77 K (see Fig. 4.11). This advantage should contribute considerably to the high-performance potential of devices based on modulation-doped Si/SiGe heterostructures like n-channel quantum-well MODFETs and MOSFETs. The experimental peak electron mobility enhancements as a function of Ge mole fraction in the virtual substrate is shown in Fig. 4.12.

### 4.1.3 Hole Mobility

For a long time, p-type Si devices have lagged behind n-type devices, both with respect to their performance characteristics as well as their understanding. Currently, strained p-MOSFETs are in the key focus in advanced logic technologies. A detailed modeling of p-type Si devices has been hampered by the highly complex band structure and scattering mechanisms that govern the dynamics of holes. For unstrained-Si at room temperature, holes occupy the top two bands: the heavy and light hole bands. With the application of strain, the hole effective mass becomes highly anisotropic due to band warping, and the energy levels become mixtures of the pure heavy, light, and split-off bands. Thus, the light and heavy hole bands lose their meaning, and holes increasingly occupy the top band at higher strain due to the energy splitting [60]. The band warping is also responsible for the fact that different types of strain (namely, the biaxial tensile and uniaxial compressive stress) behave differently.

For holes, the valence-band structure of Si is more complex than the conduction band. It is this complex band structure as well as valence band warping under strain that results in a much larger hole than electron mobility enhancement. The advances in strained-Si/SiGe technology have completely changed

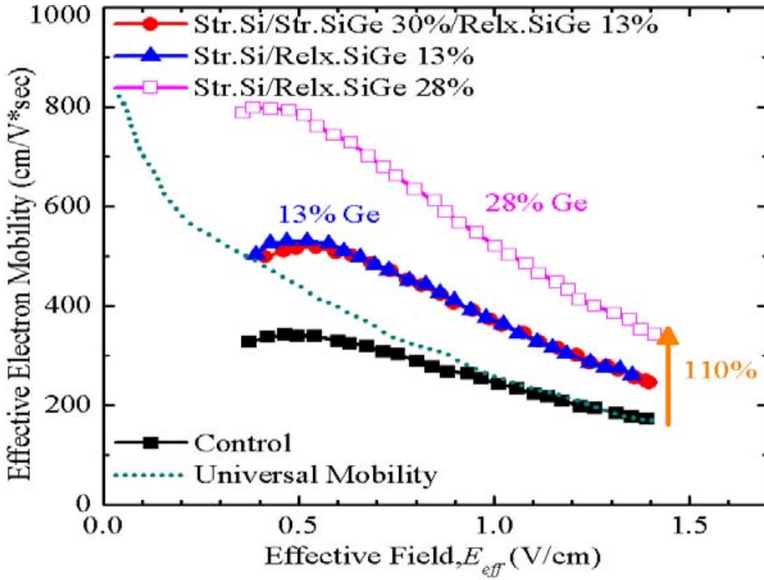



**FIGURE 4.12**
Measured electron mobility enhancement in strained-Si MOSFETs. After K. Rim et al., *IEEE VLSI Tech. Symp. Dig.*, 2002(98-99). With permission.

this situation. Theory now predicts that the characteristics of hole-type SiGe devices can match that of the electrons - an important step for continuing the miniaturization of CMOS devices. However, the understanding of hole mobility enhancement in strained-Si is more difficult due to the complex valence band structure of Si. Oberhuber et al. [64] have reported the theoretical prediction of the hole mobility enhancement in strained-Si using k-p formalism. The experimental peak hole mobility enhancements as a function of Ge mole fraction in the virtual substrate is shown in Fig. 4.13. Hole mobility enhancement decreases with increasing vertical field.

Nakatsuji et al. [65] have computed the quantum confinement of two-dimensional (2D) hole gas in inversion layer of strained-Si p-MOSFETs directly from a pseudopotential band structure [66]. The authors presented a detailed analysis of the main features of hole subband structure, i.e., the splitting energy and the effective mass, which are significantly affected by the strain. The basis of their study was the nonlocal empirical pseudopotential calculations with spin-orbit interactions for unstrained and strained-Si based on the work of Rieger and Vogl [67].

Fig. 4.14 shows the measured (symbols) and the calculated (lines) effective mass for inversion layer holes in unstrained and strained-Si MOSFETs as a function of sheet hole density. The classical cyclotron effective mass data is also shown. A good agreement is obtained as shown in Fig. 4.14. The energy

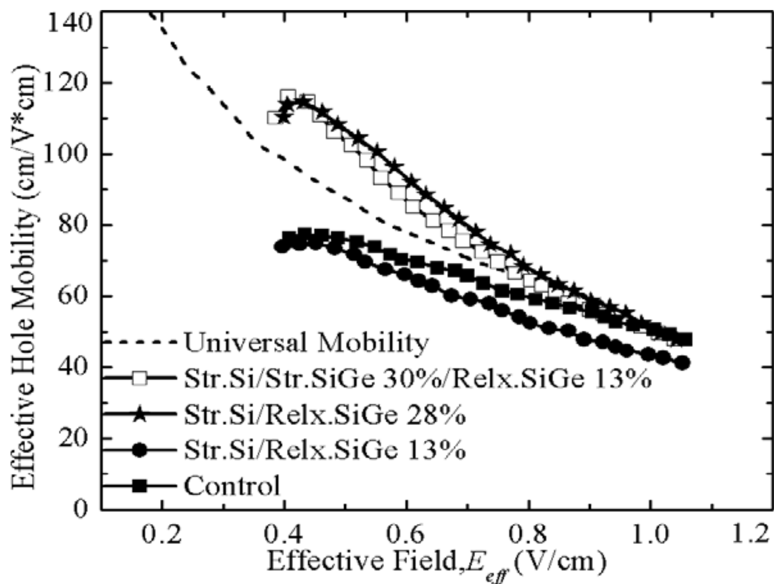



**FIGURE 4.13**
Measured hole mobility enhancement in strained-Si MOSFETs. After K. Rim et al., *IEEE VLSI Tech. Symp. Dig.*, 2002(98-99). With permission.

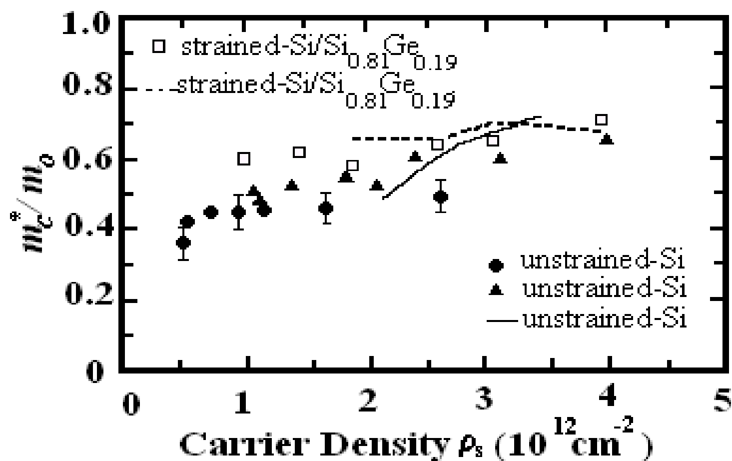



**FIGURE 4.14**
Measured (symbols) and calculated (lines) effective mass for inversion layer holes in unstrained and strained-Si as a function of sheet hole density. After H. Nakatsuji et al., *IEEE IEDM Tech. Dig.*, 2002(727-730). With permission.

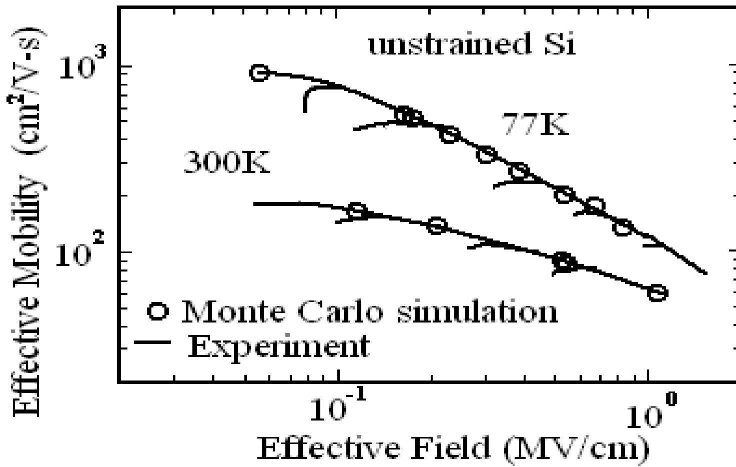



**FIGURE 4.15**
Calculated hole mobility in unstrained-Si p-MOSFET as a function of the effective electric field. After H. Nakatsuji et al., *IEEE IEDM Tech. Dig.*, 2002(727-730). With permission.

range considered for hole is very low (< 10 meV), and the effective masses for holes in unstrained and strained-Si are almost the same. However, the density of states mass in high energy regime (> 26 meV at 300 K) is reduced by the strain. On the other hand, the effective mass in the lowest subband increases with surface electric field and thus the mobility enhancement in high field region is smaller than that in the low-field region.

In order to model the surface roughness (SR), Pirovano's [68] model is employed, and the screening of the SR potential was evaluated using the screening parameter given by Fetter [69]. The physical parameters were adjusted to yield the experimental universal mobility [70] at 300 K and 77 K. A good agreement was obtained as is shown in Fig. 4.15, which confirms the validity of the scattering models.

In Fig. 4.16, a theoretical prediction for the mobility enhancement in strained-Si is shown. In the simulation, no scattering parameters except for the band structure change due to the presence of strain was included. A mobility enhancement was observed in strained-Si only for Ge concentration larger than 20%. Fig. 4.17 shows the transient velocity overshoot of 2D holes in strained-Si p-MOSFETs. The larger overshoot was observed in strained-Si because of the suppressed interband scattering.

### 4.1.4 Field Dependence

The transverse field dependence of MOS device parameters has assumed a greater importance because the thinner gate dielectrics and higher doping

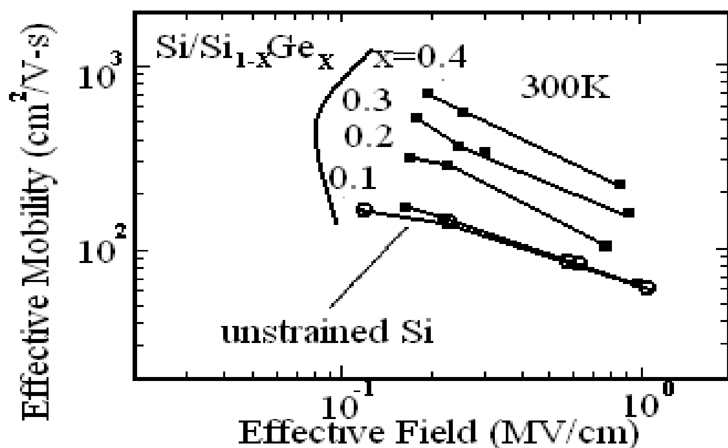



**FIGURE 4.16**
Calculated hole mobility in strained-Si grown on a relaxed-$Si_{1-x}Ge_x$ substrate at 300 K. The scattering rate parameters used are the calibrated ones for unstrained-Si. After H. Nakatsuji et al., *IEEE IEDM Tech. Dig.*, 2002(727-730). With permission.

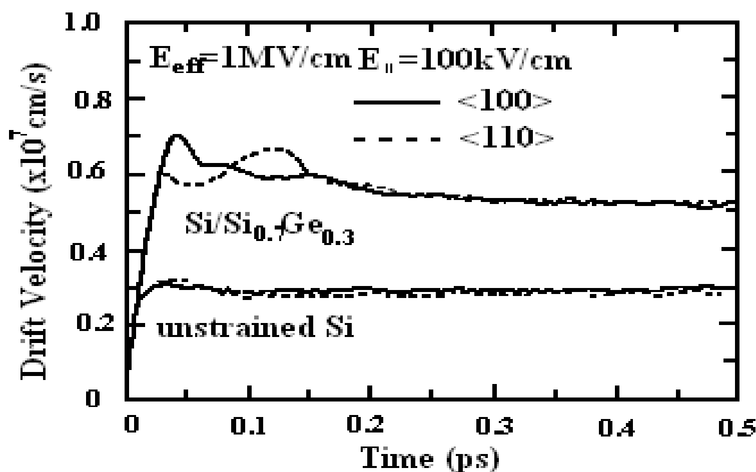



**FIGURE 4.17**
Transient velocity overshoot with a sudden application of a 100 kV/cm field parallel to the <100> and <10> directions at 300 K for unstrained-Si and strained-Si on a relaxed-$Si_{0.7}Ge_{0.3}$ substrate. After H. Nakatsuji et al., *IEEE IEDM Tech. Dig.*, 2002(727-730). With permission.

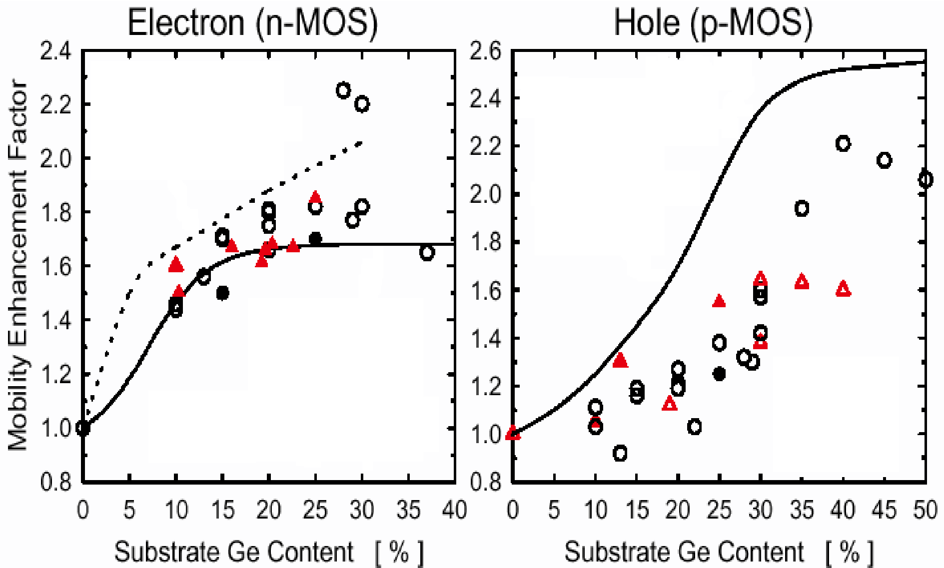



**FIGURE 4.18**

Experimental and theoretical enhancement factors of electron and hole mobility, defined by the ratio of mobility in strained-Si to that in unstrained-Si MOSFETs, as a function of mobility. Since the enhancement factor of hole mobility is dependent on $E_{eff}$, the maximum values, typically seen in low $E_{eff}$ region, are plotted in this figure. After S.-I. Takagi, ULIS 2004 Lecture Materials.

levels used in submicron MOSFETs lead to very high transverse electric fields well above 0.5 MV/cm. It is well known that such high fields cause a degradation in device performance. The variation of the effective mobility with electric field is often used as a basis of comparison of MOS devices developed and for computer aided design. In order to investigate the carrier mobility enhancement in the strained-Si MOSFETs, the effective mobility as a function of effective vertical electric field was calculated for several devices. In the following, we examine the electron and hole mobility enhancement at high vertical field in strained-Si.

Fig. 4.18 summarizes the reported experimental and enhancement factors as a function of Ge content of electron and hole mobility, defined by the ratio of mobility in strained-Si to that in unstrained-Si MOSFETs. Here, since the enhancement factor of hole mobility is dependent on $E_{eff}$, as described below, the maximum values, typically seen in low $E_{eff}$ region, are plotted in this figure. The increased mobility results in an increase in channel drive current for a given device design, leading to improved performance. The difference lies in the amount of germanium that is required. SiGe layers with

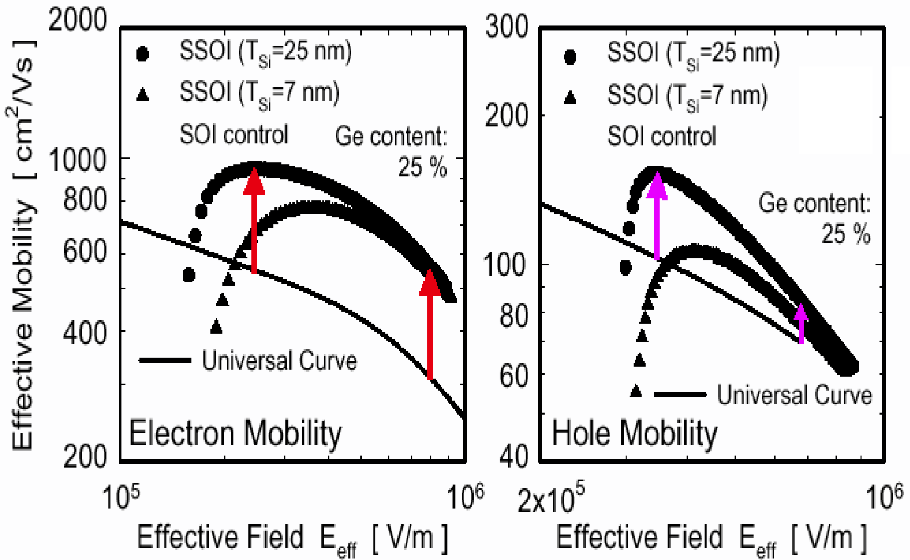



**FIGURE 4.19**
Typical $E_{eff}$ dependence of electron and hole mobility of strained-Si MOSFETs. After S.-I. Takagi, ULIS 2004 Lecture Materials.

Ge concentrations of $\sim$ 25-30% increase channel drive current up to 35% for n-channel devices, but do little to boost p-channel performance. The hole mobility is only improved if the Ge concentration is increased to 40%. In part, this is because the Ge is not electrically active in n-channel strained-Si devices, but it can be in p-channel. The problem with higher levels of Ge is that it requires more complicated buffer layer structure to reduce defects.

Fig. 4.19 shows the typical $E_{eff}$ dependence of electron and hole mobility [22]. The enhancement factor for electrons is known to be almost constant, irrespective of $E_{eff}$ [29, 71]. The electron mobility enhancement in the moderate $E_{eff}$ region can be explained well by the decrease in phonon scattering [72], as seen in agreement with theoretical results in Fig. 4.18. On the other hand, the origin of the enhancement in high $E_{eff}$ region, where surface roughness scattering dominates, is still unclear [71], though there exists a model that roughness can be reduced in strained-Si [73]. In contrast, the enhancement factor for hole mobility decreases with an increase in $E_{eff}$. A possible explanation on this phenomenon is that the increase in the subband energy due to the quantum confinement can compensate the subband splitting energy between the light hole and heavy hole band, leading to smaller subband energy difference in higher $E_{eff}$ [22, 74, 75]. While the recent theoretical calculations [28] seem to be almost consistent with this interpretation, further quantitative examination is needed.

It is experimentally and theoretically confirmed that mobility of roughly

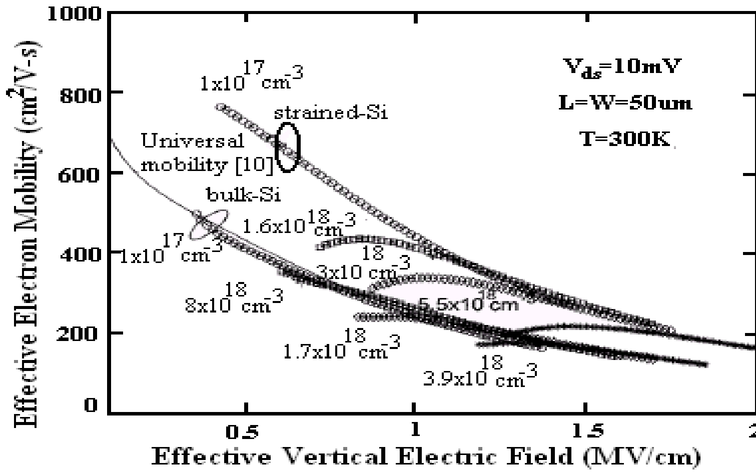



**FIGURE 4.20**
Effective electron mobility vs. vertical effective electric field, $E_{eff}$, for various channel doping concentrations for unstrained- and strained-Si n-MOSFETs. After H. M. Nayfeh et al., *IEEE Electron Dev. Lett.*, Vol. 24, 2003(248-250). With permission.

twice as high as in conventional MOSFETs can be obtained for strained-Si n- and p-MOSFETs. As for n-MOSFETs, the experimental results are in fairly good agreement with the theoretical calculation [71, 72]. On the other hand, the experimental results in p-MOSFETs are smaller than the theoretical predictions [64, 65] and Ge contents higher than in the theories are needed to obtain the same amount of the enhancement. This discrepancy in the hole mobility enhancement between the experiments and the theories demands further studies on transport properties of holes in strained-Si.

### 4.1.5 Doping Dependence

The strain dependence of electron inversion layer mobility has been measured in n-MOSFETs fabricated on a strained-Si/relaxed-$Si_{1-x}Ge_x$ heterostructure [20, 76]. For a Ge content of 20% in the underlying relaxed-$Si_{1-x}Ge_x$, the electron mobility of the strained-Si is enhanced by roughly 75%. Inversion layer mobility measurements in strained-Si n-MOSFETs fabricated using a typical MOSFET process including high temperature steps and with various channel doping concentrations, achieved by boron ion implantation, are compared with coprocessed bulk-Si n-MOSFETs.

Short-channel effects in a MOSFET are usually controlled by increasing the channel doping concentration, thus suppressing the spread of source and drain depletion layers. Therefore, as the channel length is scaled down to 0.1 $\mu$m and below, the doping concentration must be raised to values around $1\times10^{18}$

$cm^{-3}$. As the strained-Si shows a mobility enhancement in the range of 1.7-2× in relatively lightly doped channels [19, 77], it is important to study the dependence of electron mobility on channel doping. Also, such high impurity concentrations lead to an increase in threshold voltage and to a severe decrease in electron mobility, which strongly degrades the electrical properties of the device [78].

Researchers from MIT [79] have studied the inversion layer mobility measurements in strained-Si n-MOSFETs fabricated using a typical MOSFET process, including high temperature steps and with various channel doping concentrations, achieved by boron ion implantation, and have compared with coprocessed bulk-Si n-MOSFETs. It is found that a near-universal mobility relationship with vertical effective electric field, $E_{eff}$, exists for strained-Si and bulk-Si n-MOSFETs for all channel implant doses in their study. Significant mobility enhancement for $E_{eff}$ up to 2 MV/cm (1.5-1.7×) is obtained for channel doping concentrations ranging from $10^{17}$-$6\times10^{18}$ $cm^{-3}$.

For low field, the mobility for strained-Si devices decreases toward the unstrained-Si data due to Coulomb scattering by channel dopants. At low inversion layer charge, and high doping where Coulomb scattering is important, the mobility enhancement for the strained-Si devices is reduced. This is likely due to Coulomb scattering by channel dopants, agreeing with theoretical predictions that the enhancement of strained-Si electron mobility over unstrained-Si is reduced when Coulomb scattering is the dominant carrier scattering mechanism [80].

Fig. 4.20 shows that at high inversion layer charge, the strained-Si mobility data plotted vs. effective electric field display a universal behavior independent of doping with enhancement of 1.5-1.7× over unstrained-Si. The unstrained-Si data also show universal behavior and agree closely with reported data [70]. For strained and unstrained devices at high $E_{eff}$, where phonon and surface-roughness are the dominant scattering mechanisms for electron transport, mobility was calculated by fitting the analytical universal mobility expressions using Matthiessen's rule with channel doping concentrations $3.9\times10^{17}$ $cm^{-3}$ and $5.5\times10^{17}$ $cm^{-3}$. Good agreement of calculated and measured data for both cases has been found. The fitting parameters used for unstrained Si closely follow those published by Liang et al. [81]. The ratio is 1.75, equal to the observed mobility enhancement.

## 4.2 Carrier Lifetime

Carrier generation lifetime ($t_g$) is an important material parameter and is used to characterize epi layers. Carrier generation lifetime ($t_g$) in strained-Si/SiGe has been investigated using capacitance transient method in MOS structure.

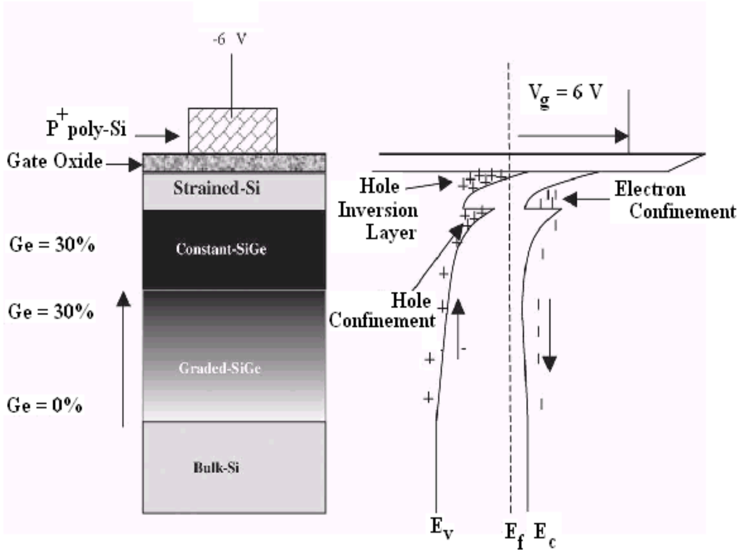



**FIGURE 4.21**
Schematic cross-section of the MOS test structure on strained-Si/SiGe and associated energy band line-up under a negative gate bias. After L. K. Bera et al., *Appl. Surf. Sci.*, Vol. 224, 2004(278-282). With permission.

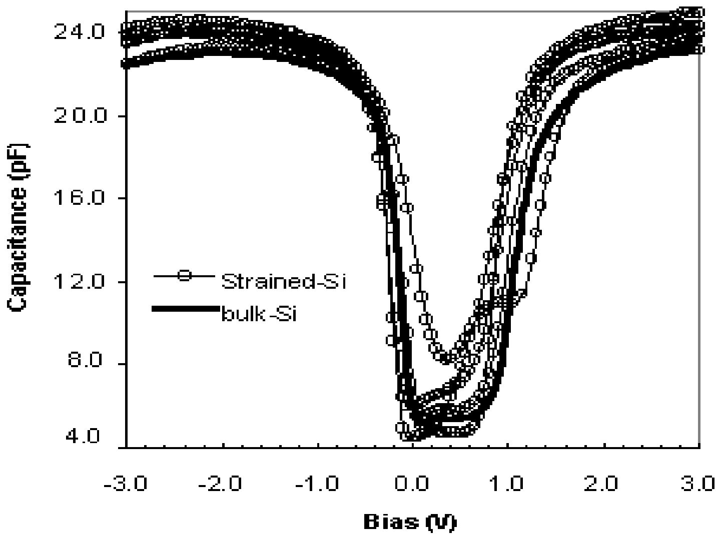



**FIGURE 4.22**
Quasistatic C-V characteristics of MOS capacitors on strained-Si and bulk-Si wafers. After L. K. Bera et al., *Appl. Surf. Sci.*, Vol. 224, 2004(278-282). With permission.

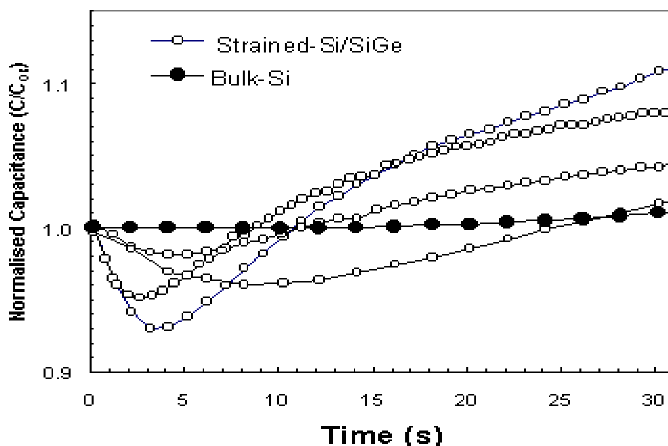



**FIGURE 4.23**
Normalized capacitance (C/$C_{ox}$) vs. time characteristics of strained-Si/SiGe and bulk-Si samples. All open circle data correspond to strained Si of different thickness. After L. K. Bera et al., *Appl. Surf. Sci.*, Vol. 224, 2004(278-282). With permission.

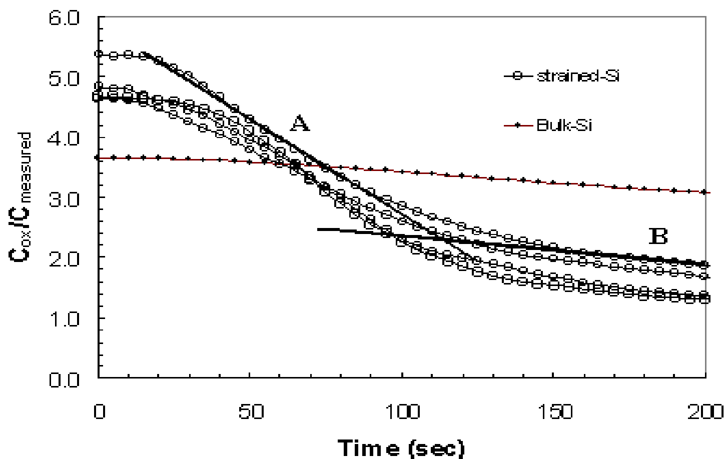



**FIGURE 4.24**
$C_{ox}/C_{measured}$ vs. time characteristics of strained-Si/SiGe and bulk-Si samples. All open circle data correspond to strained-Si of different thickness. After L. K. Bera et al., *Appl. Surf. Sci.*, Vol. 224, 2004(278-282). With permission.

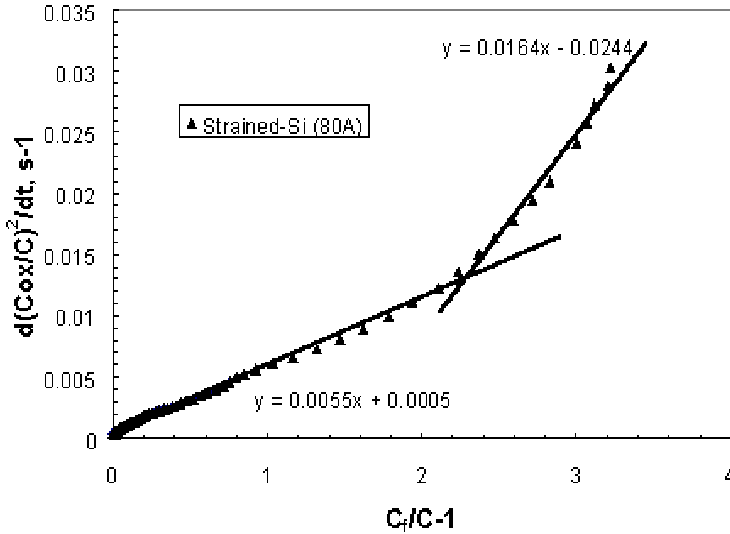



**FIGURE 4.25**
Zerbst plot from transient capacitance data of (80 Å) strained-Si/SiGe sample. After L. K. Bera et al., *Appl. Surf. Sci.*, Vol. 224, 2004(278-282). With permission.

Generation lifetime in strained-Si and SiGe buffer layer is estimated from the segments of Zerbst plot having different slopes. Fig. 4.21 shows the schematic cross-section of the MOS test structure and the associated band lineup under a negative gate bias. Fig. 4.22 shows the quasistatic C-V characteristics of MOS capacitors grown on different strained-Si and control-Si wafers. The oxide thickness (57 Å) was estimated from the accumulation capacitance. The C-T measurements are performed using HP-4284A LCR meter. The MOS structure is first biased to accumulation by applying a voltage of $> 6$ V. At time $t = 0$, it is set to deep depletion by fast switching to 6 V. At this moment, the capacitance reaches the lowest value. It takes a while to form the inversion layer by carrier generation. The time for reaching the equilibrium inversion capacitance is therefore proportional to the generation lifetime.

Fig. 4.23 shows the transient response of different strained-Si and bulk-Si samples normalized to the inversion capacitance at time $t = 0$. An initial decrease in capacitance on the transient characteristics appeared in all strained-Si samples and is in contrast to the bulk-Si capacitors. This can be explained using the band diagram of the capacitor on strained-Si/SiGe after the application of negative voltage (deep depletion condition) to the gate terminal as shown in Fig. 4.21. During generation of electron-hole pair in the strained-Si layer, a fraction of electrons will be confined at the strained-Si/relaxed-SiGe

interface due the large conduction band offset. However, the generated holes drift to the inversion layer at $SiO_2$/strained-Si interface. A fraction of the holes generated in the SiGe layer will be stored up in strained-Si/SiGe interface due to the valence band offset. If one assumes a net positive charge as a result of the above charge generations in the semiconductor region, it amounts to the placement of negative image charges on the gate electrode which increases the depletion width. Thus the measured capacitance decreases initially until the inversion charge in the $SiO_2$/strained-Si interface dominate over the charges in the potential wells as shown in Fig. 4.23. The extent of the capacitance excursion would depend upon the carrier generation dynamics in a complex system containing different epi layers, interfaces, and associated defects.

Fig. 4.24 shows the $C_{ox}/C_{measured}$ vs. time plot to investigate the nonlinear characteristics of capacitance transient response. Note that all the samples show different segments having different slopes. Based on the depletion width, calculated from the capacitance transient response, two sections are chosen from each curve to calculate the lifetime. The initial steep slopes (segment-A in Fig. 4.24) are attributed to the fast generation rate in SiGe layer. Segment-B shows the flattening characteristics, which indicates a slow generation rate in strained-Si layer. Typical Zerbst plot of strained-Si (80 Å) MOS capacitor is shown in Fig. 4.25. The slope of the Zerbst plot from the strained-Si segment is lower compared to the SiGe segment. This clearly indicates higher generation lifetime for strained-Si compared to the SiGe layer, as expected. Generation lifetime calculated from the slope in SiGe layer ranges from 20 to 90 ms. The generation lifetime in strained-Si and coprocessed bulk-Si is 120-170 ms and 177 ms, respectively. Comparable values of lifetime in both strained-Si and coprocessed bulk-Si indicate that the "electrical" quality of strained-Si film is approaching that of the bulk. The smaller bandgap of strained-Si can also contribute to lower carrier lifetime compared to bulk-Si.

## 4.3 Mobility: Thickness Dependence

In the following, we discuss the effects of strained-Si layer thickness on mobility with silicon thicknesses below and above the critical thickness. The strained-Si critical thickness ($t_{crit}$) is defined as the maximum strained-Si film thickness for which a dislocation-free silicon film is thermodynamically stable [82]. Fiorenza et al. [83] have demonstrated that exceeding the critical thickness does indeed have a deleterious effect on the MOSFET behavior, although the mobility enhancement is not greatly affected by exceeding the critical thickness (see Fig. 4.26). However, MOSFETs fabricated on strained-Si films greater than the critical thickness show an increased level of off-current.

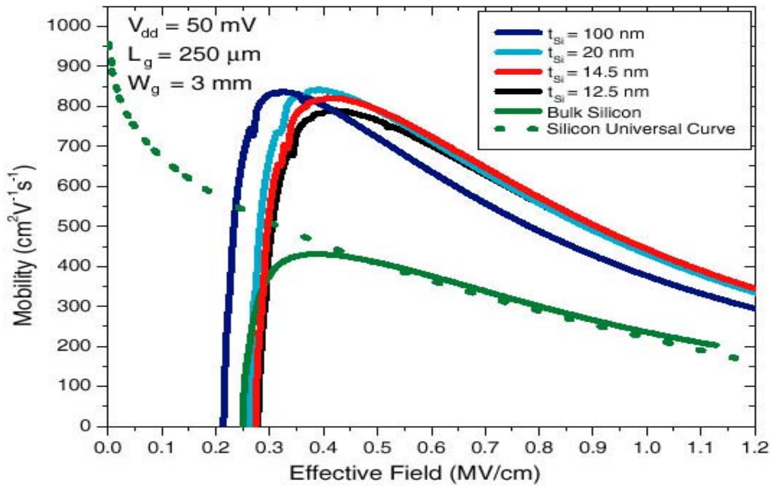



**FIGURE 4.26**
Electron mobility of a bulk-Si MOSFET and of strained-Si MOSFETs with different strained silicon thickness. The strained-Si MOSFETs show enhanced mobility even with silicon films significantly greater than the strained-Si critical thickness. After J. G. Fiorenza et al., *Semicond. Sci. Technol.*, Vol. 19, 2004(L4-L8).

The effects of misfit dislocations at the silicon/SiGe interface were studied by comparing the characteristics of n-MOSFETs fabricated on the four different wafers with the same graded-SiGe buffer construction but with four different strained-Si thicknesses and the bulk-Si control wafer. The substrates were grown by LPCVD on p-type Si wafers. The SiGe graded layer was grown at a grading rate of 10% Ge $\mu m^{-1}$ to a final Ge content of 20%, yielding a 2 $\mu$m-thick graded layer. The graded layer was capped with a 2 $\mu$m uniform 20% SiGe layer. The SiGe buffer layer growth temperature was 1000$^o$C. After removal of the cross-hatch using chemical mechanical polishing, an additional 1.5 $\mu$m layer of uniform 20% SiGe was grown and then the strained-Si layer was grown. The strained-Si layer growth temperature was 650$^o$C. The four different strained-Si thicknesses were 12.5 nm, 14.5 nm, 20 nm and 100 nm. n-MOSFETs with gate lengths between 0.5 $\mu$m and 200 $\mu$m were fabricated on these four substrates and on a bulk-Si Control wafer. The fabrication process used a single mask level and featured n-type doped polysilicon gates, 5-nm thick thermally grown $SiO_2$ gate oxide, $SiO_2$ spacers, and titanium salicide source gate/drain metallization. The significant components of the device fabrication thermal budget were the gate oxide growth (800$^o$C, 45 min), dopant activation (1000$^o$C, 1 s), silicide formation anneal (675$^o$C, 5 min) and sinter (450$^o$C, 30 min).

Low-field electron mobility was extracted from large-area transistors (L = 250 $\mu$m and W = 3 $\mu$m) on each of the four strained-Si substrates and the

silicon control wafer using the split C-V technique. The low-field mobility was enhanced by a factor of about 1.8 relative to the control bulk-Si wafer. This is somewhat unexpected, as the thicker films are partially relaxed due to the introduction of misfit dislocation and, therefore, is expected to exhibit less electron mobility enhancement. Thus, strain and mobility enhancement are largely maintained even for silicon film thicknesses far greater than $t_{crit}$. However, the motivation for maintaining strained-Si film thicknesses below $t_{crit}$ is not just to avoid the loss of strain and the loss of mobility enhancement, but that of high leakage current [83].

In general, mobility enhancement in strained-Si is observed due both to the decrease in effective mass, and the intervalley scattering of electrons or interband scattering of holes. Several workers have also reported a reduction in the carrier mobility in strained-Si [22, 84]. The strain in the silicon layer causes the six-fold degenerate valleys of the silicon conduction band minimum to split into two groups: two lowered valleys with the longitudinal effective mass axis perpendicular to the interface, and four raised valleys with the longitudinal mass axis parallel to the interface. The combination of a lighter effective mass and reduced intervalley scattering gives rise to higher electron mobility [85], and also electron velocity overshoot [58].

Recent studies have shown the feasibility of growth of high-quality strained-Si layers on SiGe-on-insulator where a high-quality crystalline relaxed-SiGe layer is deposited with a thickness of less than 10-nm on $SiO_2$ substrates by separation-by-implanted-oxygen (SIMOX) techniques or by solid phase epitaxy (SPE). Using strained-Si layers on SiGe-on-insulator (SiGe-OI) structures and growing a thin strained-Si layer on the relaxed SiGe-OI substrate, both n- and p-channel strained-Si/SiGe-OI MOSFETs have been fabricated.

The strained-Si/SiGe-OI structures combine the advantages of strained-Si/SiGe-bulk MOSFETs and those of ultrathin fully-depleted SOI MOSFETs. In Si/SiGe-OI structures, advantages of SOI and strain are combined with the use of an insulating substrate: the resulting structure provides good control of short-channel effects, lower parasitic capacitance, and higher radiation tolerance, as well as mobility values that are much higher than those found in conventional SOI MOSFETs. Both the electron and hole mobilities at low electric field are found to decrease in thinner strained-Si structures, and reduction of hole mobility enhancement has also been reported at high electric field [8]. A small reduction in the electron mobility with thin strained-Si layers is attributable to the quantum mechanical confinement effect of the inversion layer electrons, originating in the conduction band offset of the strained-Si layers.

The hole mobility at a low field depends on the strained-Si thickness. The diffusion of Ge atoms into the $SiO_2$/strained-Si interface is also found to generate interface states near the valence band edge, leading to the reduction in hole mobility in the lower $E_{eff}$ region through Coulomb scattering. Moreover, the decrease in hole mobility enhancement in both thin and thick strained-Si structures at the higher electric field is caused by the reduction

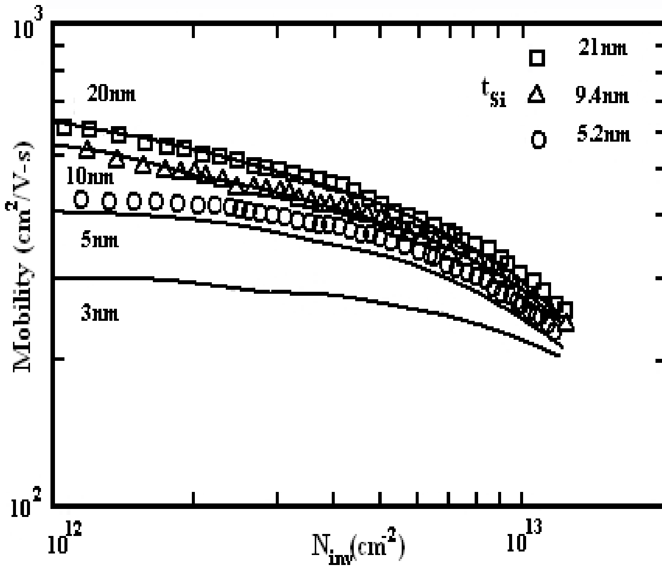



**FIGURE 4.27**
Comparison of the mobility calculated from spatial distribution of carriers with recent experimental results. Source: MIT Annual Report 2002.

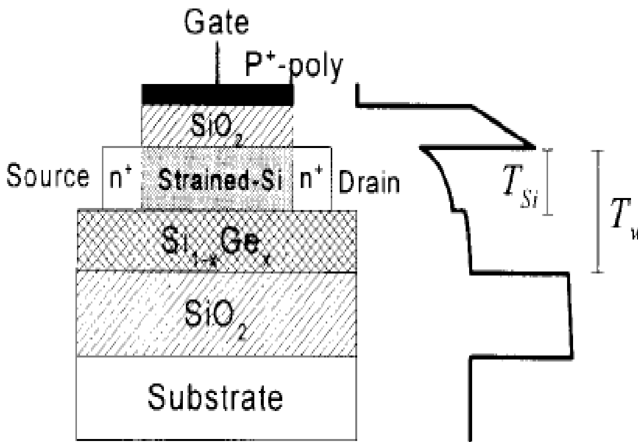



**FIGURE 4.28**
Typical strained-Si MOSFET structure on SiGe-OI substrate. Si layer thickness $T_{Si}$, SiGe-layer thickness $T_{SiGe}$, gate oxide thickness $t_{ox}$ 55 nm, buried oxide thickness $T_{box}$ 580 nm. After F. Gamiz et al., *Appl. Phys. Lett.*, Vol. 80, 2002(4160-4162). With permission.

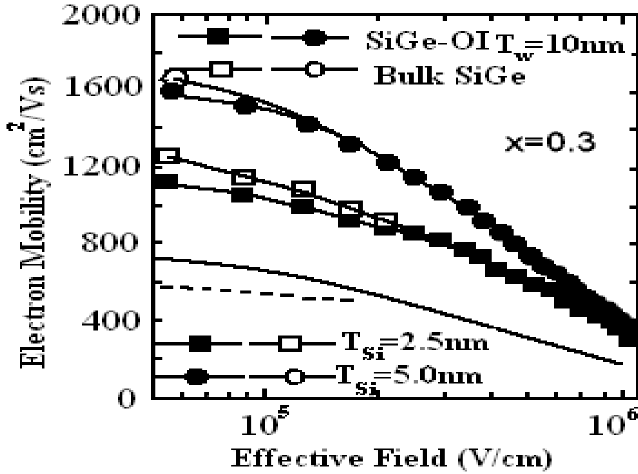



**FIGURE 4.29**
Electron mobility curves vs. the transverse effective field for the simulated device of Fig. 4.28 at room temperature for different values of $T_{Si}$ ($T_w = T_{Si} + T_{SiGe}$ = 10 nm, x = 0.3). After F. Gamiz et al., *Appl. Phys. Lett.*, Vol. 80, 2002(4160-4162). With permission.

of the energy splitting between the heavy and the light hole bands, with an increase in the electric field. In order to realize high performance strained-Si CMOS in the sub-100 nm region, the total thickness of the strained-Si and the relaxed-SiGe layers must be thin enough to suppress the short-channel effects without increasing the channel dopant impurity concentration. Therefore, it is important to understand the mechanisms responsible for the reduction in the carrier mobility in strained-Si and formulation of device design guidelines for strained-Si CMOS, involving the Ge content of the SiGe layers and the strained-Si thickness.

Nevertheless, to gain these advantages, also to provide a low off-state leakage current, to allow operation at low voltages, and to avoid the floating body effect, the Si/SiGe structure sandwiched between the two oxides must be thin enough ($T_w = T_{Si} + T_{SiGe} \leq 30$ nm) where $T_{Si}$ and $T_{SiGe}$ are the thicknesses of the strained-Si layer and of the relaxed-SiGe layer, respectively. On the other hand, in order to maintain the strain, the silicon layer must be sufficiently thinner than the SiGe layer. Therefore, taking into account both conditions, ($T_w \leq 30$ nm and $T_{Si} \leq T_{SiGe}$), the strained-Si layer thickness has to be reduced to a very low value ($T_{Si} \leq 10$ nm) and it shows the existence of a lower limit to the $T_{Si}$ that may be used in ultrathin Si/SiGe-OI MOSFETs, if one wants to keep a high mobility value.

In order to suppress short-channel effects, strained-Si film thickness should

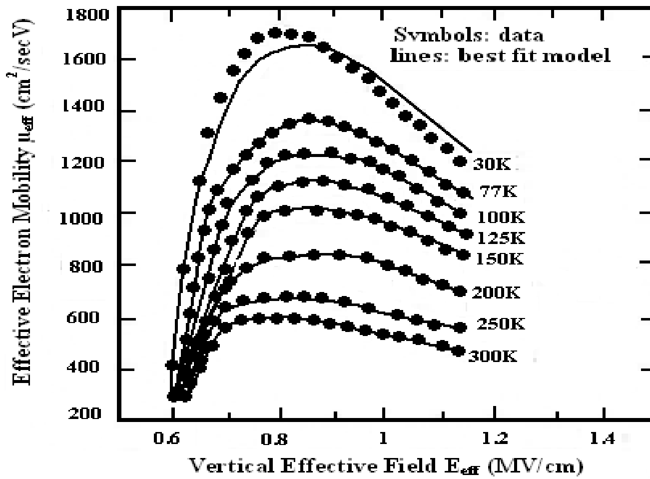



**FIGURE 4.30**
Comparison of measured strained-Si n-MOSFET mobility. Symbols and solid lines represent the experimental data and the best fit model, respectively. Source: MIT 2001 Annual Report.

be scaled down along with shrinking the gate length (e.g., 3-nm thick for 10-nm gate length). It has been shown both theoretically and experimentally that the mobility is degraded for films thinner than 10 nm. To consider this degradation in device simulation and optimization, researchers from MIT have developed a method to calculate the mobility in ultrathin SOI films based on the spatial distribution of the carriers. Self-consistent Schrodinger-Poisson simulations are used to obtain carrier distribution in the film. Experimental universal mobility is then used to correlate this spatial distribution with the mobility. As shown in Fig. 4.27, results of these calculations are in good agreement with recent experimental data.

Gamiz et al. [86] have shown by Monte Carlo (MC) simulation that electron mobility is greater when strained-Si inversion layers are grown on SiGe-on-insulator and is strongly dependent on the strained-Si layer thickness, $T_{Si}$. This is due to an increase of the phonon scattering rate that occurs as the silicon layer thickness is reduced as consequence of the greater confinement of the carriers due to the discontinuity of the conduction band at the Si/SiGe interface (see Fig. 4.28), which, however, partially counteracts the increase in the mobility achieved by the strain. This effect is less important as the Ge mole fraction, x, is reduced, and as the value of $T_{Si}$ increases. This effect is less important as the germanium mole fraction, x, is reduced, and as the value of $T_{Si}$ increases. For $T_{Si} \geq 20$ nm, mobility does not depend on $T_{Si}$, and the maximum values of the mobility are obtained. In strained-Si/SiGe-OI structures both high electron and hole mobilities with a germanium mole fraction as low as 0.1 have been observed. From Fig. 4.29, the following

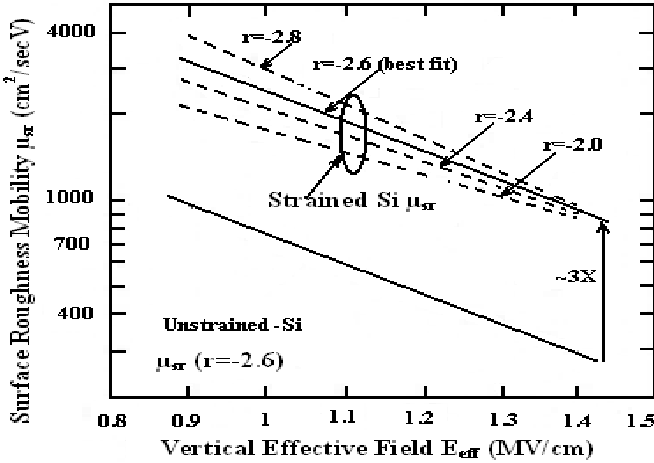



**FIGURE 4.31**
Extracted surface roughness limited mobility for strained- and unstrained-Si devices. For strained-Si, the fit to the data is down with various values of the parameter, r, which is the power dependence of this mobility term on the vertical effective field (r = 2.6±0.2 gives the best fit to the data). Source: MIT Annual Report 2001.

observations are made:

(a) in case of the strained case, there is an increase in the electron mobility compared to the unstrained case, as has been experimentally observed, due to the combination of a lighter conduction effective mass and smaller intervalley scattering rate,

(b) the mobility curves corresponding to strained-Si/SiGeOI MOSFETs are slightly lower than those corresponding to strained-Si/SiGeOI bulk MOSFETs, especially at low transverse effective fields as shown in Fig. 4.29. This is due to the greater phonon-scattering rate in the SiGe-OI case, arising from the greater confinement of the carriers. An increase in the energy difference between the subbands, which causes the reduction in the intervalley scattering rate and a decrease in the conduction effective mass contribute to an increase in the electron mobility, and

(c) the electron mobility in the strained-Si/SiGe-OI MOSFETs strongly depends on the $T_{Si}$ value; the mobility curve corresponding to $T_{Si}$ = 5 nm is much higher than that corresponding to $T_{Si}$ = 2.5 nm. This phenomenon is not exclusive of the Si/SiGe-OI structures, but it also appears in bulk-Si/SiGe structures. The reason for this is the increase in the phonon-scattering rate. The discontinuity in the conduction band minimum just at the Si/SiGe interface means the electrons are more closely confined near the interface.

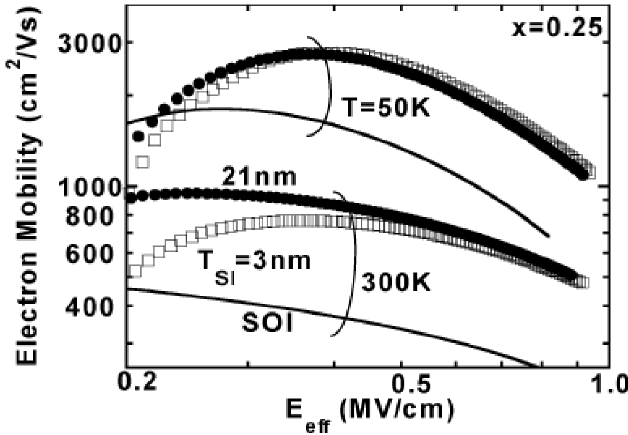



**FIGURE 4.32**
Electron mobility characteristics of thick (circles) and thin (squares) strained-Si structures at various temperatures. The solid lines show the control-SOI data. The electron mobility of the thin (3 nm) strained-Si structures is degraded in the low-field region at room temperature compared to thick (21 nm) structures. However, this degradation disappears at low temperatures (50 K). After T. Mizuno et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(1114-1121). With permission.

## 4.4 Mobility: Temperature Dependence

One method of understanding the mechanism of mobility enhancement is to study the vertical effective field dependence of the mobility as a function of temperature. For high vertical fields (> 1 MV/cm), it is anticipated that surface roughness scattering will dominate transport. The impact of strain on surface roughness scattering is not clearly understood yet. Lowering the measurement temperature reduces the phonon component of scattering, and thus enables one to observe the remaining mobility terms, which include Coulomb and surface roughness scattering, in strained- and unstrained-Si devices.

Measurements and extraction of the mobility as a function of temperature from strained-Si n-MOSFETs (20% Ge) and unstrained-Si control wafers with similar doping profiles are shown in Fig. 4.30. The surface roughness mobility term dominates the mobility at low temperatures (e.g., 30 K) and at room temperature at high vertical effective fields. The surface roughness mobility term appears to be enhanced by almost a factor of three, at a vertical effective field of 1.4 MV/cm. The results suggest that surface-specific mechanisms (see Fig. 4.31) may be involved in the strain-induced electron mobility

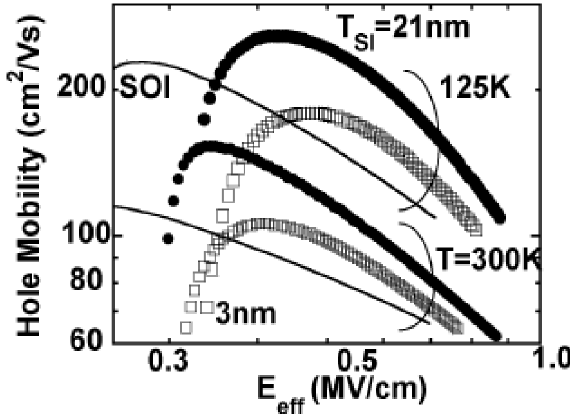



**FIGURE 4.33**
Hole mobility characteristics of 21-nm (closed circles) and 3-nm (open circles) thick strained-Si structures at various temperatures. The solid lines show the data of control-SOIs. The hole mobility of the thin strained-Si structures is degraded at both the room- and low-temperature (125 K) conditions. Moreover, the hole mobility enhancement decreases with increasing electric field. However, this degradation at high field becomes smaller at low temperatures. After T. Mizuno et al., *IEEE Trans Electron Dev.*, Vol. 51, 2004(1114-1121). With permission.

enhancement.

In order to evaluate the scattering mechanism for inversion-layer mobility, electron mobility properties of thin and thick strained-Si structures have been studied at various temperatures [75]. Fig. 4.32 shows the electron mobility behaviors of control-SOIs (solid line), and thin (squares) and thick strained-SOIs (circles) at room (300 K) and low temperatures (50 K), as a function of $E_{eff}$. The decrease in electron mobility in the thinner strained-Si layer is observed under room temperature conditions and low $E_{eff}$, while the electron mobility enhancement is still observed in high even at room temperature. However, this mobility reduction low disappears at 50 K.

Fig. 4.33 shows the hole mobility characteristics in control- (solid line) and strained-SOIs with thick (circles) and thin (squares) strained-Si layers at 300 and 125 K. One of the hole mobility characteristics in strained-SOI p-MOSFETs is that the hole mobility enhancement factor decreases with increasing $E_{eff}$ at both temperatures. One possible reason of this decrease in the hole mobility enhancement at higher $E_{eff}$, is the decrease or the disappearance of the energy splitting between the light hole and the heavy hole bands associated with an increase in $E_{eff}$, because the effective mass of the heavy hole bands is heavier than that of the light holes.

## 4.5 Diffusion in Strained-Si

Stress effects on dopant diffusion/activation become more important as transistor dimensions are scaled into the nanoscale regime. Stresses exist in devices due to various process steps like STI or silicidation. Further, high levels of stress are being induced purposefully to enhance carrier mobility. Consequently, these result in asymmetric nonhydrostatic stress states in silicon. A detailed understanding of the expected complex interaction of stress on dopant diffusion and activation is lacking. Density functional theory (DFT) based calculations are used to predict stress effects on the formation and migration of point defects and dopant/defect.

Strained-Si on a relaxed-SiGe buffer layer on (001) Si is typically grown at low temperatures. To be fully incorporated into conventional Si technology, however, these layers have to withstand further processing, i.e., annealing and ion implantation. It has been demonstrated that strained-Si relaxes via misfit and threading dislocation propagation and Ge interdiffusion [87, 88, 89] after high-temperature post annealing. To fully understand the stability of these layers, the influence of ion implantation and subsequent activation annealing must be determined. Ion implantation can create point defects, which can serve as nucleation sites for relaxation-induced defects or aid in the interdiffusion of Ge; thus, relaxation can be accelerated.

In the near future, strain engineering will be extensively used to enhance CMOS device performance. The extent to which strain engineering can be successfully combined with conventional ion implantation doping of ultrashallow source-drain regions will depend on the effects of strain on dopant diffusion and activation at high dopant concentrations. Experiments on the thermal equilibrium diffusion of dopants incorporated epitaxially at low concentrations into strained-Si layers have shown significant strain effects, arising from changes in the formation energies of diffusion for the mobile diffusing species.

In order to accurately predict the dopant profiles after thermal annealing, it is important to understand the dopant diffusion in strained-Si. An understanding of n-type dopant diffusion in strained-Si and SiGe, specifically the formation of the source/drain regions in the nMOS and the n-body region in the pMOS, is essential for the success of strained-Si technology. Strained-Si and SiGe is usually grown on Si substrates and depending on Ge concentration and layer thickness, the layers can be strained or relaxed. Various dopants react differently to the amount and type (compressive/tensile) of strain present in the SiGe layer. In the past, studies have been performed on Sb and B diffusion in relaxed-SiGe and also on their diffusion behavior under different types of strain [90, 91]. Arsenic diffusion in relaxed-SiGe has been studied in the past using the radiotracer method [92] as well as ion implantation [93].

It has been shown that tensile strain enhances the stability of the As-interstitial defects mediating diffusion, band of the extended interstitial de-

fects driving transient enhanced diffusion, which leads to a complex strain-dependent diffusion behavior. In addition, n-type dopant in SiGe technology is expected to be important for the development of other devices such as thermoelectric generators, novel MOSFET structures that utilize SiGe source/drain contacts, and also SiGe HBT technology.

Eguchi et al. [94] compared the diffusivity of ion implanted As and Phosphorus (P) in SiGe and Si under equilibrium conditions. Fig. 4.34 shows the structure of the samples. Sample wafers were ion implanted As (30 keV, $4\times10^{14}$) or P (30 keV, $8\times10^{14}$). The samples were then furnace annealed at from 950$^o$C to 1050$^o$C for 30 minutes in an Ar ambient. The oxide layers were stripped and samples were analyzed by Secondary Ion Mass Spectrometry (SIMS). The SIMS data was compared with TSURPREM-4 simulations. For comparison, Si control pieces were processed and analyzed at the same time as the SiGe samples.

Fig. 4.35 shows the comparison of As and P effective diffusivity in the temperature range from 950$^o$C to 1050$^o$C under equilibrium conditions. The diffusivities of n-type dopant are observed to be enhanced in SiGe compared to Si, in contrast to the diffusion of B, which is known to be retarded in SiGe relative to B diffusion in Si. The expression for the concentration-dependent diffusivity of P in Si must be multiplied by a factor of roughly two, in order to match the P in SiGe profile. And the effective arsenic diffusivity enhancement factor is roughly 7 for diffusion in SiGe compared to diffusion in Si. The absolute value of As effective diffusivity is still 3 times lower than that of phosphorus. These results suggest that As is the better choice for shallow source/drain than P.

As diffusion behavior under transient diffusion conditions has also been studied. Fig. 4.36 shows the comparison of As profiles in Si and SiGe after the short-time annealing (1000$^o$C 1, 20, 60 sec). Fig. 4.37 shows the time dependence of As effective diffusivities in Si and SiGe. TRD (Transient Retarded Diffusion) is observed for As diffusion in SiGe, while TED (Transient Enhanced Diffusion) is observed at As diffusion in Si. This TRD shows the possibility of forming As ion implanted source/drains in SiGe as shallow as those in Si by optimizing the annealing conditions.

## 4.6 Process-induced Strained-Si

Longitudinal uniaxial compressive stress is observed to cause much larger hole mobility enhancement at low and moderate stress on p-MOSFETs compared to biaxial tensile strain. In this section, we explain why the uniaxial hole mobility enhancement is present at high vertical electric field. Mobility enhancement in the case of uniaxial strain has been investigated by Thompson

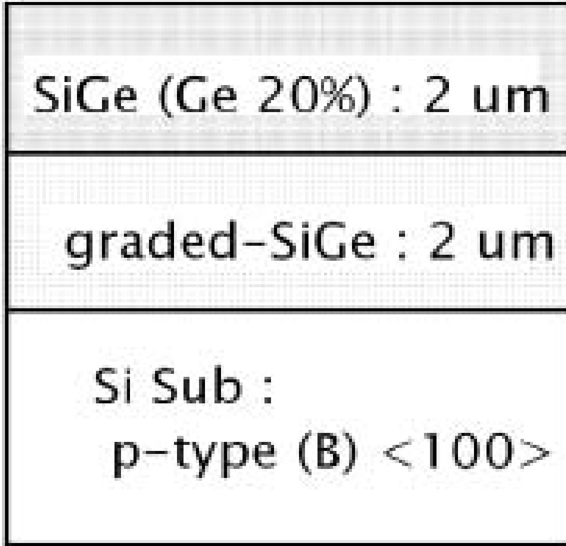



**FIGURE 4.34**
Structure of SiGe samples. Source: MIT Annual Report 2001.

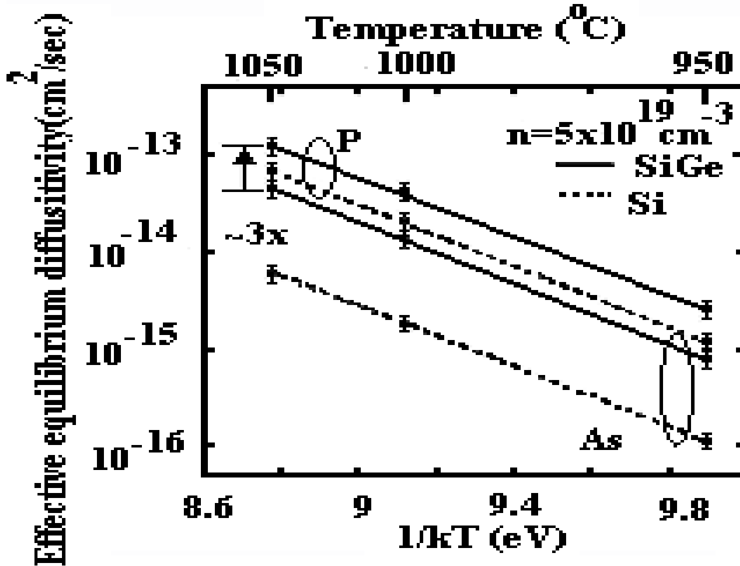



**FIGURE 4.35**
Comparison of As and P effective diffusivities in Si and SiGe under equilibrium conditions. Source: MIT Annual Report 2001.

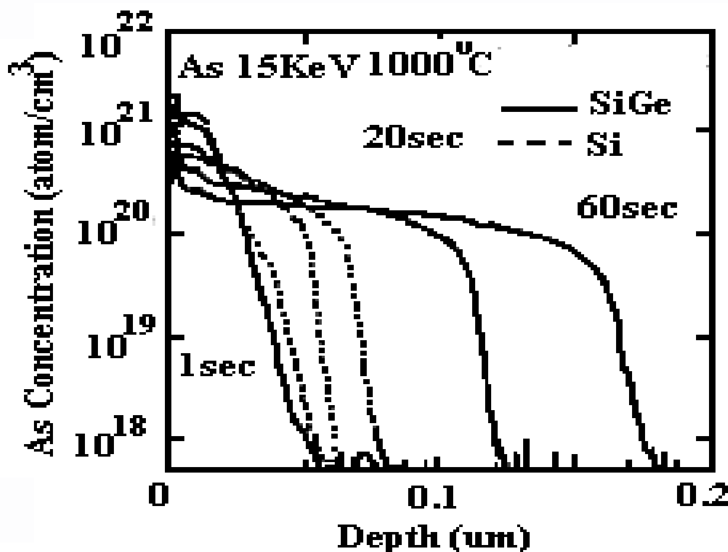



**FIGURE 4.36**
Comparison of As profiles in Si and SiGe under transient diffusion conditions. Source: MIT Annual Report 2001.

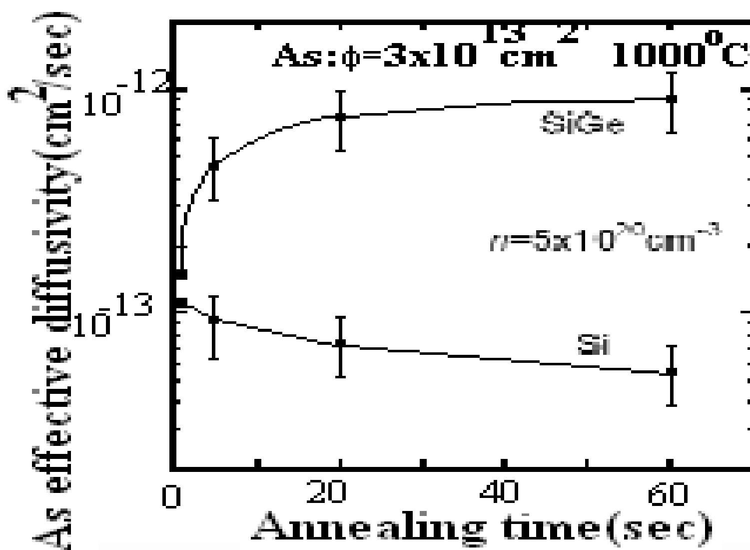



**FIGURE 4.37**
Time dependence of As effective diffusivities in Si and SiGe under transient diffusion. Source: MIT Annual Report 2001.

et al. [11]. Uniaxial process-induced strain is generally applied either parallel (longitudinal) or perpendicular (transverse) to the direction of MOSFET current flow. In the case of process-induced strain, such as in a CMOS logic technology, small strain is desirable since low strain creates less defects and requires minimum variation in the process technology. This also helps in keeping lower thermal cycles which help in avoiding the strain relaxation. It is known that in small strain case, the piezoresistance varies linearly with strain (typically with 1% nonlinearity at 250 MPa) [95] and the piezoresistance coefficients can be used as a guide as to which strain maximizes the mobility enhancement.

Using the piezoresistance coefficients one can quantify the strained-Si mobility enhancement. In the case of conventional Si wafers with (001) surface and wafer notch on the [110] axis, for uniaxial process-induced strain applied either in parallel (longitudinal) or perpendicular (transverse) to the direction of MOSFET current flow, aligned to the <100> axes, the mobility can be expressed as follows [11]:

$$\frac{\Delta\mu}{\mu} \approx \mid \pi_{\|}\sigma_{\|} + \pi_{\perp}\sigma_{\perp} \mid \tag{4.16}$$

where the subscripts $\|$ and $\perp$ refer to the directions parallel and transverse to the current flow in the plane of the MOSFET, $\Delta\mu/\mu$ is the fractional change in mobility, $\sigma_{\|}$ and $\sigma_{\perp}$ are the longitudinal and transverse stresses, and $\pi_{\|}$ and $\pi_{\perp}$ are the piezoresistance coefficients expressed in $\text{Pa}^{-1}$. $\pi_{\|}$ and $\pi_{\perp}$ can be expressed in terms of the three fundamental cubic piezoresistance coefficients $\pi_{11}$, $\pi_{12}$, and $\pi_{44}$. The longitudinal and transverse piezoresistance coefficients, first measured by Smith [96], are given in Table 4.1. The piezoresistance coefficients should take into account the two-dimensional (2D) nature of transport in MOSFETs and depend on temperature and doping [97]. However, for simplicity, Thompson et al. [11] have computed the $\pi_{\|}$ and $\pi_{\perp}$ using only the bulk coefficients neglecting these effects. The calculated values are shown in Table 4.1. It has been reported that the bulk coefficients quantitatively match the strained-Si experimental data and are consistent with the work reported by Steegen et al. [98].

From Table 4.1, the large positive value of $\pi_{\|}$ for p-MOSFETs with <110> channel orientation shows significantly less strain is needed for enhanced hole mobility when introduced with longitudinal compression vs. in-plane biaxial strain. This is quantified by evaluating Eqn. 4.16 and using

$$E^{o} = \frac{stress}{strain} = \frac{\sigma_{\|}}{\frac{\Delta l}{l}} \tag{4.17}$$

where $E^{o}$ is the Young's modulus, and $(\Delta l/l)$ is the strain. The evaluation shows only 0.2% lattice displacement is required for a 30% mobility gain compared to the biaxial case where > 1% displacement is required. It has been shown that the uniaxial strain increases the hole mobility for the 45-nm gate length transistor by 50%.

**TABLE 4.1**
Longitudinal and transverse piezoresistance coefficients evaluated for standard layout and wafer orientation (units of $10^{-12}$ cm$^2$dyne$^{-1}$). After S. E. Thompson et al., *IEEE Electron Dev. Lett.*, Vol. 25, 2004(191-193). With permission.

| | <100> | <100> | <110> | <110> |
|---|---|---|---|---|
| Polarity | $\pi_{\parallel}$ | $\pi_{\perp}$ | $\pi_{\parallel}$ | $\pi_{\perp}$ |
| N or P | $\pi_{11}$ | $\pi_{12}$ | $(\pi_{11} + \pi_{12} + \pi_{44})/2$ | $(\pi_{11} + \pi_{12} - \pi_{44})/2$ |
| N-type | -102 | 53.4 | -31.6 | -17.6 |
| P-type | 6.6 | -1.1 | 71.8 | -66.3 |

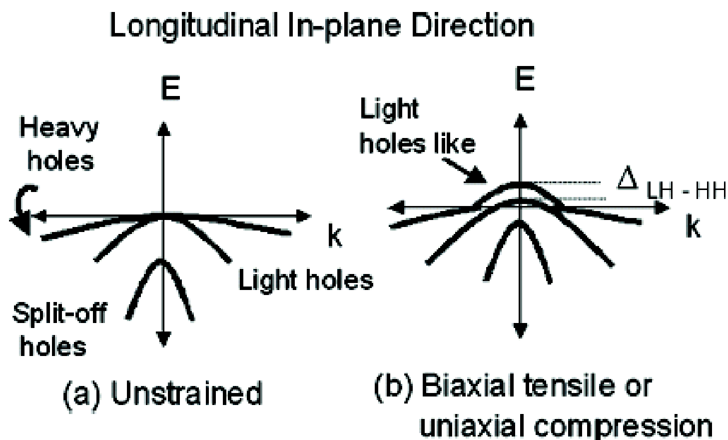



**FIGURE 4.38**
Simplified hole valance band structure for longitudinal in-plane direction. (a) Unstrained and (b) strained-Si. After S. E. Thompson et al., *IEEE Electron Dev. Lett.*, Vol. 25, 2004(191-193). With permission.

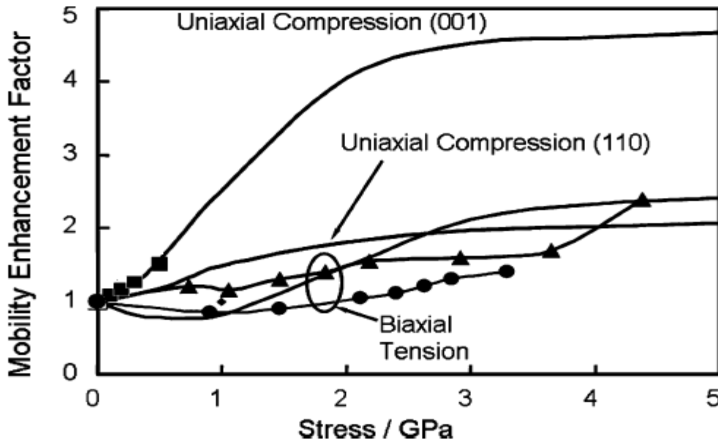



**FIGURE 4.39**
Calculated and experimental data for longitudinal compressive and biaxial tensile stress enhanced mobility vs. stress. Note that the maximum predicted Si inversion layer hole mobility enhancement is estimated to be $\sim$ 4 times higher for uniaxial stress on (100) wafer and $\sim$ 2 times higher for biaxial stress on (100) wafer and for uniaxial stress on a (110) wafer. After S. E. Thompson et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2006(1010-1020). With permission.

For strained-Si, the explanation for the hole mobility enhancement has been offered by several workers. For both the biaxial tensile and uniaxial compressive stress, a key component of the enhanced mobility results from the lowest energy subband having a low conductivity effective mass [28, 64]. To explain the field dependence of the mobility in biaxial tensile strain, Fischetti et al. [28] showed that reduction in hole mobility enhancement with vertical field results from the quantization associated with the confining surface potential of the MOSFET. The strain-induced formation of low in-plane effective mass was first calculated by Bir and Pikus [99]. Fig. 4.38 summarizes the hole band structure for both the unstrained- and strained-Si. The valance bands are plotted for the in-plane direction of the MOSFET channel. Both uniaxial and biaxial strain lift the degeneracy in the valance band and cause the bands to not only shift, but to change shape as shown in Fig. 4.38. An important observation is that for the process-induced strain, the hole mobility enhancement is present even at a high vertical field. This result is consistent with other uniaxial strained-Si where the mobility enhancement is also observed at a high field: mechanical wafer bending [100] and nitride capping layer [5, 6].

Process-induced strain-enhanced mobility can be predicted from the piezoresistance coefficients. Since piezoresistance is not expected to vary linearly with stress level above $\sim$ 250-500 MPa, for higher levels of stress, caution in us-

ing piezoresistance coefficients is needed. When evaluating the various stress options, it is also important to comprehend the maximum possible mobility enhancement. Thompson et al. [101] have used a set of scattering parameters that fit the experimental data for hole mobility enhancement under biaxial tensile stress (see Fig. 4.39). The calculations include acoustic and optical phonon and surface roughness scattering. This set of scattering parameters shows that the dominant mechanism responsible for biaxial tensile-stress mobility enhancement (at large stress) is reduced optical phonon scattering. Acoustic phonon scattering is only slightly altered due to the changes in the density of states. Using the same scattering parameters, mobility enhancement for uniaxial stress on (001) and (110) wafers was calculated, as shown in Fig. 4.39, and compared to uniaxial stress data from literature. The model shows good agreement. The maximum predicted Si inversion-layer hole mobility enhancement is estimated to be $\sim$ 4 times higher for uniaxial stress on (100) wafer and $\sim$ 2 times higher for biaxial stress on (100) wafer and for uniaxial stress on a (110) wafer. The larger maximum mobility enhancement on a (001) wafer results from the high density of states in the top band.

### 4.6.1 Hole Mobility

Hole mobility enhancement have been studied by several authors for uniaxial stress [40, 55, 102]. A novel computationally efficient model for stress-modulated hole mobility, suitable for continuum transport simulators, has been developed by Obradovic et al. [103]. The physics based model considers band structure modulation due to stress, and can predict the mobility behavior over a wide range of stress, electric fields, and current directions. In the model, a simplified heavy-hole valence bandstructure is considered by a pair of intersecting ellipsoids in k-space. Eqn. 4.18 specifies the mobility enhancement as a function of relative populations (f), current direction and ($\theta$), carrier temperature T, and the curvature masses and is given by:

$$\mu_{enh} = 2\frac{m_t m_l}{m_t + m_l}\left[cos^2\theta\left(\frac{f_1}{m_{l1}} + \frac{f_2}{m_{l2}}\right) + sin^2\theta\left(\frac{f_1}{m_{l1}} + \frac{f_2}{m_{l1}}\right)\right] \tag{4.18}$$

$$f_w = \frac{1}{1 + exp(-\frac{\Delta}{kT})} \tag{4.19}$$

$$f_h = \frac{1}{1 + exp(\frac{\Delta}{kT})} \tag{4.20}$$

Eqns. 4.19 and 4.20 define the components of the in-plane stress tensor: shear (s), biaxial (b), and asymmetric (a). Uniaxial stress along the direction of current flow lowers the energy of the W regions relative to the H region (see Fig. 4.40(a)). This results in a repopulation of holes into the W region, which is more favorable for conduction than the H region due to the larger gradient

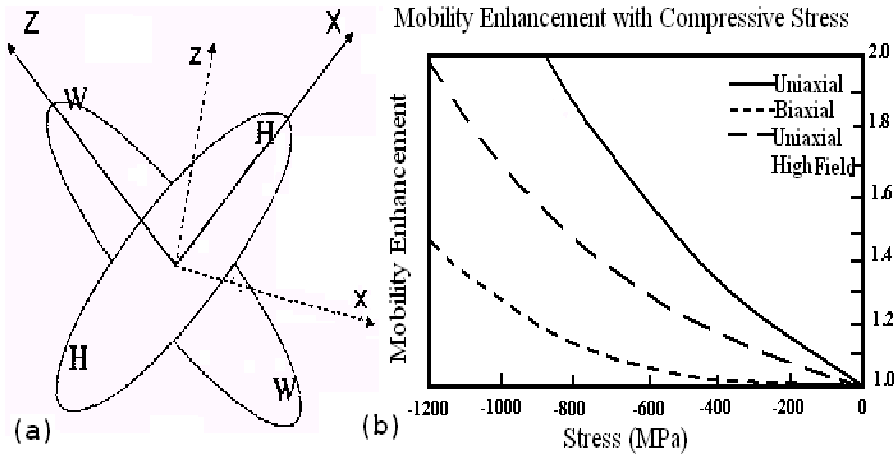



**FIGURE 4.40**
(a) Simplified heavy-hole valence band. (b) The modeled stress response for compressive stresses is shown. The behavior is strongly nonlinear at high stress values in sharp contrast to the Piezoresistive model. After B. Obradovic et al., *IWCE Tech. Dig.*, 2004(26-27). With permission.

in k-space along the current flow direction. Thus, mobility is enhanced. Conversely, a longitudinal tensile or transverse compressive stress results in an H to W repopulation, and a corresponding reduction of mobility.

The analytic model is constructed by parametrizing the relative energy shift of the two ellipsoids and their curvatures as functions of the in-plane stress tensor, expressed in terms of the crystal coordinates of the Si unit cell. The relative fractional populations $f_w$ and $f_H$ of the W and H ellipsoids are calculated, based on the stress-dependent energy separation $\Delta$ and Maxwell-Boltzmann statistics (using the carrier temperature). For MOS devices with typical channel orientations (110), the perpendicular component of the current vanishes, but the mobility tensor is not isotropic under general stress conditions and field orientations. The dependence on carrier energy (or lateral field, for drift-diffusion) is introduced through the carrier temperature term in the Boltzmann statistics. This results in a reduced mobility enhancement for high-energy carriers as shown in Fig. 4.40(b).

The model has been validated and calibrated using a set of wafer-bending experiments (Fig. 4.41). Devices of various lengths (with built-in stress) are subjected to additional longitudinal or transverse stress from the wafer bending, for a total stress range (bending plus structural) of 700 MPa tensile to 800 MPa compressive. The overall agreement to data is found to be very good.

For uniaxial stress, the out-of-plane mass is calculated from the $6\times6$ Hamiltonian for Si and Ge in Fig. 4.42. Both Si and Ge show the interesting result

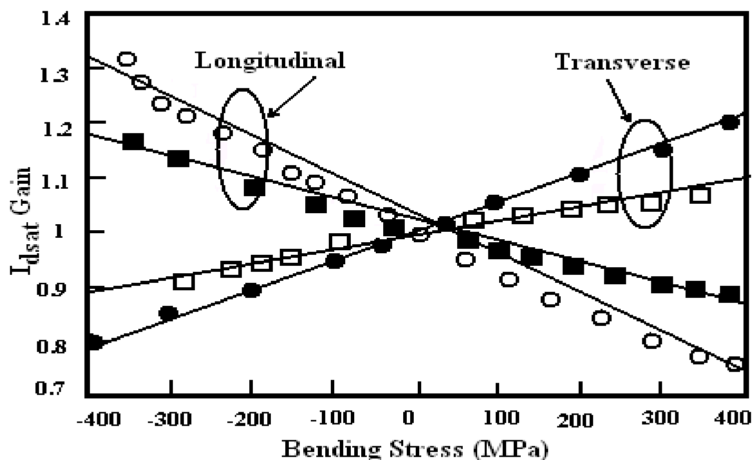



**FIGURE 4.41**
Comparison of predicted $I_{dsat}$ gain (lines) to wafer-bending data (points) for long-channel (1 $\mu$m - solid line) and short-channel devices (55 nm - dashed line). All devices have built-in stress in addition to the bending stress. After B. Obradovic et al., *IWCE Tech. Dig.*, 2004(26-27). With permission.

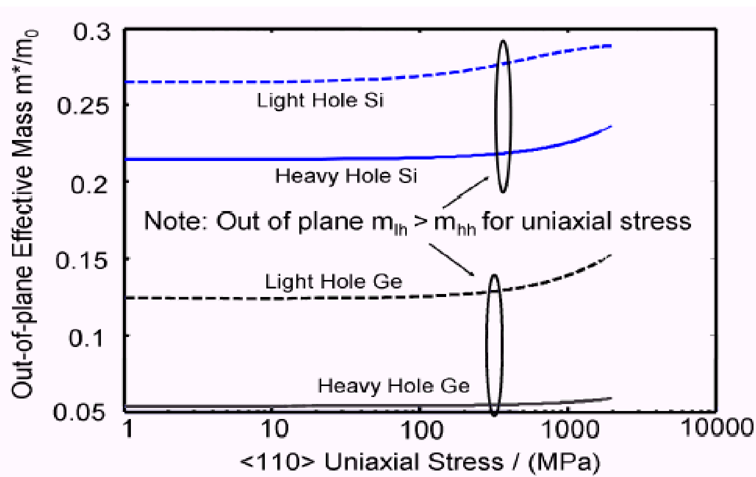



**FIGURE 4.42**
Strained-Si and Ge MOSFET out-of-plane hole effective mass vs. uniaxial strain. For both strained-Si and Ge, $m_{lh} > m_{hh}$. After S. E. Thompson et al., *IEEE IEDM Tech. Dig.*, 2004(221-224). With permission.

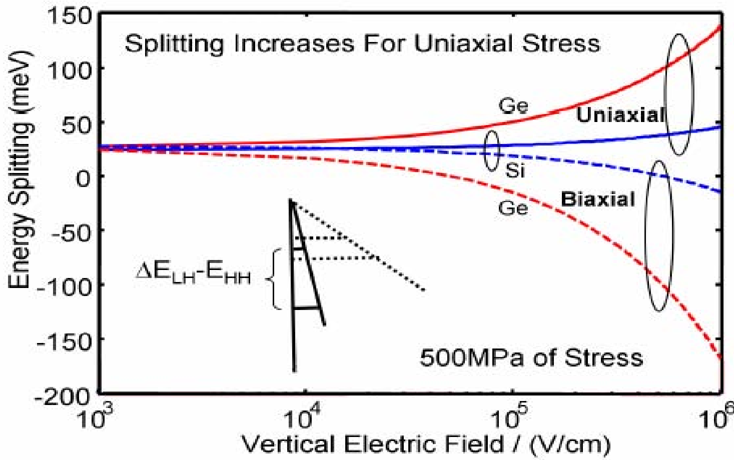



**FIGURE 4.43**
Band splitting vs. vertical electric field for 500 MPa of stress. The splitting is additive for uniaxial stress. After S. E. Thompson et al., *IEEE IEDM Tech. Dig.*, 2004(221-224). With permission.

that for uniaxial stress, $m_{lh}{}^2 > m_{hh}{}^2$ which causes the light to heavy hole band splitting to increase with surface confinement, while for biaxial stress the previously reported $m_{lh}{}^2 < m_{hh}{}^2$ causes the band splitting to be reduced by surface confinement [28]. Including confinement and for 500 MPa of uniaxial or biaxial stress, Fig. 4.43 shows the band splitting vs. vertical electric field. The band splitting caused by surface confinement is additive for uniaxial but subtractive for biaxial stress, which is consistent with the experimental hole mobility data.

In a MOSFET, the 2D surface confinement shifts the light and heavy hole bands. Using the empirical pseudopotential method, Fischetti et al. [28] have show that surface confinement cancels the biaxial stress light to heavy hole band splitting causing the enhanced mobility to be lost at high vertical fields. In contrast, uniaxial hole mobility data suggest the splitting is not canceled by surface confinement, which is a major advantage. The band splitting caused by surface confinement is additive for uniaxial but subtractive for biaxial stress, which is consistent with the experimental hole mobility data. The experimental data and band calculations support uniaxial stress having an advantageous in- and out-of-plane effective mass. The valence band warping under uniaxial and biaxial stress is shown in Fig. 4.44.

Band structures of uniaxial and biaxial stressed Si are calculated using empirical nonlocal pseudopotential method [27]. Since the conduction band edge, $\Delta E_c$, of Si is located far from the $\Gamma$ point, (0,0,0), the k·p perturbation method, which is frequently used to analyze the hole transport, is not adequate to accurately evaluate the conduction band structure in Si. The subband

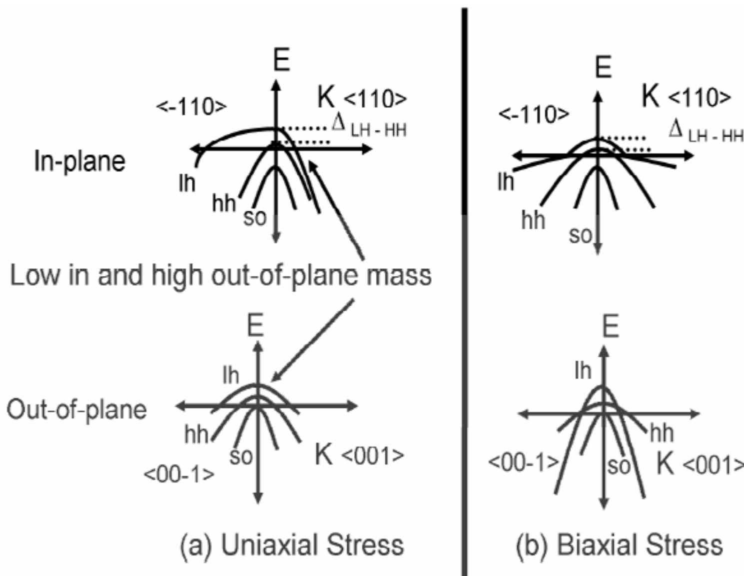



**FIGURE 4.44**
Simplified valence band E-k diagram for strained-Si. After S. E. Thompson et al., *IEEE IEDM Tech. Dig.*, 2004(221-224). With permission.

structure of the inversion layers in MOSFETs under stress is obtained using Poisson-Schrodinger solver with considering $\Delta E_c$. Then, electron mobility for relaxed and stressed MOSFETs is calculated employing phonon scattering with relaxation-time approximation. Figs. 4.45-4.47 show the calculated splits of the conduction band edge, $\Delta E_c$, under biaxial, uniaxial <110>, and uniaxial <100> stress.

### 4.6.2 Electron Mobility

The physical mechanisms of electron mobility enhancement by uniaxial stress have recently been reviewed by Thompson et al. [101]. From full-band calculations, uniaxial-stress-induced split of conduction band edge, $\Delta E_c$, and effective mass change, $\Delta m^*$, are quantitatively evaluated. The authors have experimentally and theoretically demonstrated that the energy surface of two-fold valleys in Si (001) FETs is warped due to uniaxial <110> stress, resulting in lighter $m_t$ of two-fold valleys parallel to the stress. By using calculated $\Delta E_c$ and $\Delta m^*$, experimental electron mobility enhancement is accurately modeled for biaxial, uniaxial <100>, and uniaxial <110> stress.

Figs. 4.48-4.49 show the electron mobility enhancement induced by stress, $\Delta\mu_e/\mu_e$, as a function of effective field, $E_{eff}$, under biaxial and uniaxial <100> stress. The good agreement between experimental results and calcu-

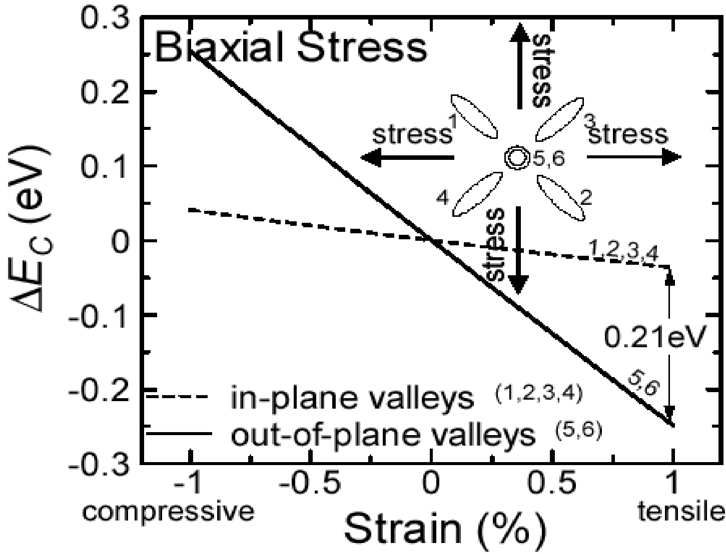



**FIGURE 4.45**
Split of conduction band edge, $\Delta E_c$, as a function of strain under biaxial stress. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

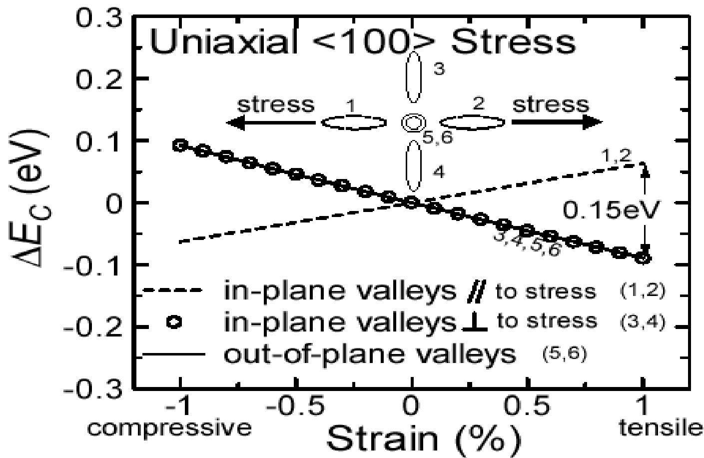



**FIGURE 4.46**
Split of conduction band edge, $\Delta E_c$, as a function of strain under uniaxial <100> stress. The strain is measured in the stressed direction. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

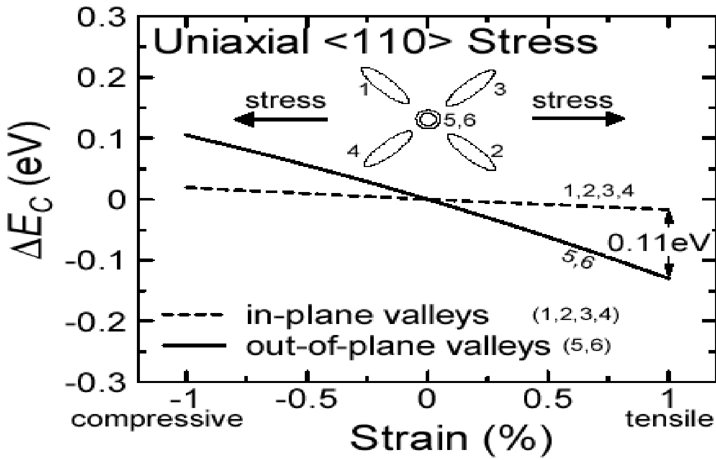



**FIGURE 4.47**
Split of conduction band edge, $\Delta E_c$, as a function of strain under uniaxial <110> stress. The strain is measured in the stressed direction. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

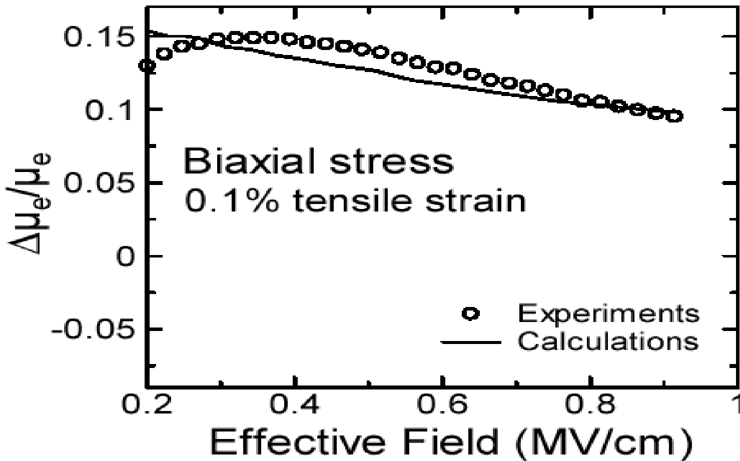



**FIGURE 4.48**
Electron mobility enhancement, $\Delta\mu_e/\mu_e$, induced by biaxial stress as a function of effective field at the tensile strain of 0.1%. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

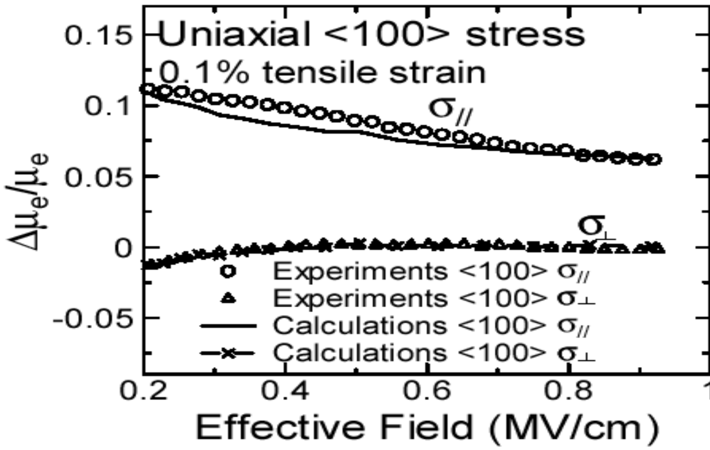



**FIGURE 4.49**
Electron mobility enhancement, $\Delta\mu_e/\mu_e$, induced by uniaxial <100> stress as a function of effective field at the tensile strain of 0.1%. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

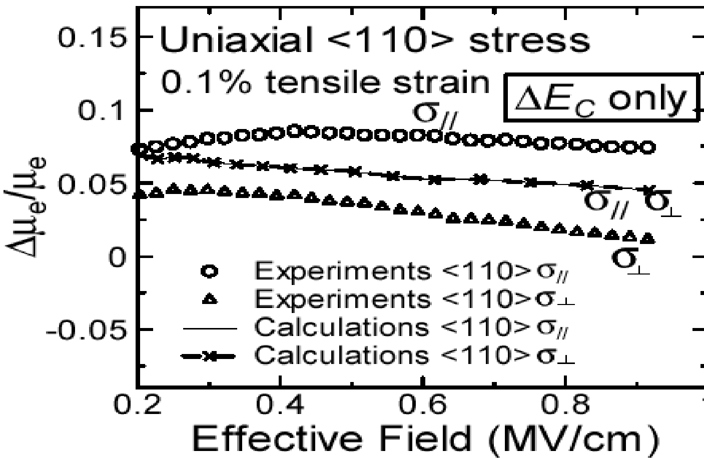



**FIGURE 4.50**
Electron mobility enhancement, $\Delta\mu_e/\mu_e$, induced by uniaxial <110> stress as a function of effective field at the tensile strain of 0.1%. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

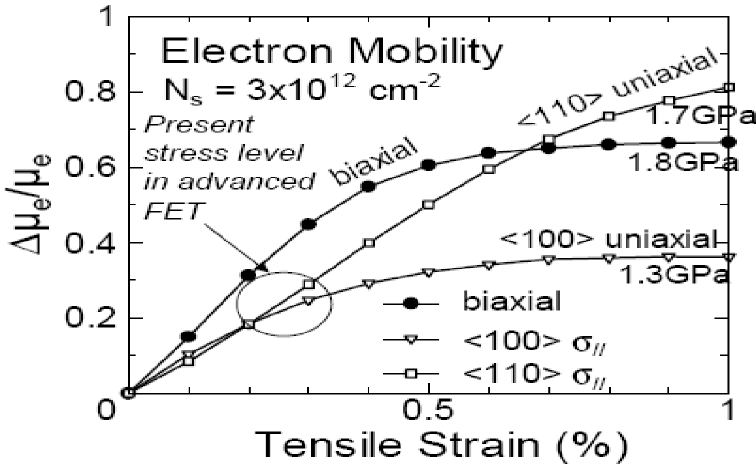



**FIGURE 4.51**
Electron mobility enhancement as a function of strain. Stress realizing 1% strain for each stress condition is described. At high stress of ~ 1.5 GPa, <110> uniaxial stress is better than biaxial stress due to m* change. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

lated ones confirms that the electron mobility enhancement is quantitatively ascribed to $\Delta E_c$ for both biaxial and uniaxial <100> stress cases. In uniaxial <110> case, however, the calculations do not reproduce the experimental electron mobility enhancements (Fig. 4.50), indicating that $\Delta E_c$ is not enough to account for the electron mobility enhancement by uniaxial <110> stress.

In order to anticipate the limit of the electron mobility enhancement by the stress engineering, electron mobility enhancements [104] by biaxial, uniaxial <100>, and uniaxial <110> stress are calculated in a wide range of strain by taking into account $\Delta E_c$ and $\Delta m^*$ (see Fig. 4.51). It has been demonstrated that electron mobility enhancement under uniaxial <110> stress does not saturate even at the stress of 1.7 GPa because of the reduction of $m_t$ at higher stress. On the other hand, electron mobility enhancements by biaxial and uniaxial <100> stress saturate at higher stress, since the change of the electron population and the suppression of the intervalley phonon scattering saturate at larger $\Delta E_c$. Thus, the uniaxial <110> stress is more advantageous than biaxial and uniaxial <100> stress in terms of electron mobility enhancement particularly at larger strain (higher stress). Uniaxial <110> stressed MOSFETs with 0.5% strain (0.85 GPa) results in a 27% higher $I_{d,sat}$. Figs. 4.52-4.54 show the $\mu_e$ enhancement, $\Delta\mu_e/\mu_e$, vs. strain characteristics of bulk FETs under biaxial, uniaxial <100>, and uniaxial <110> stress, respectively. Biaxial tensile stress shows the highest $\mu_e$ enhancement, followed by <100> $\sigma_{\|}$ and <110> $\sigma_{\|}$. Only <100> $\sigma_{\perp}$ reduces $\mu_e$.

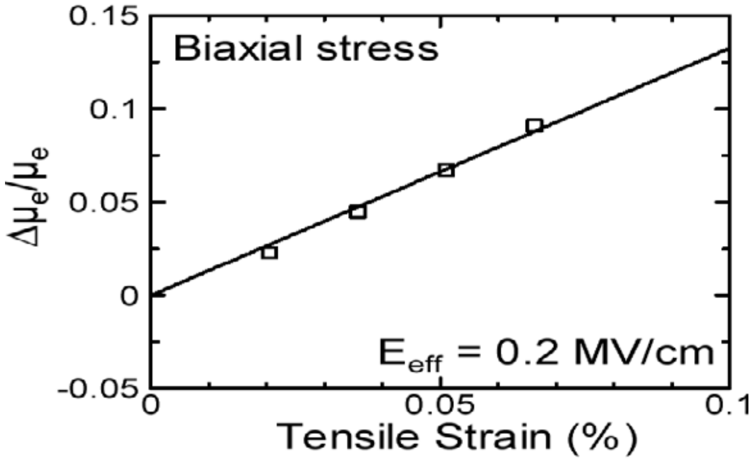



**FIGURE 4.52**
Electron mobility enhancement, $\Delta\mu_e/\mu_e$, induced by biaxial tensile stress as a function of strain. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

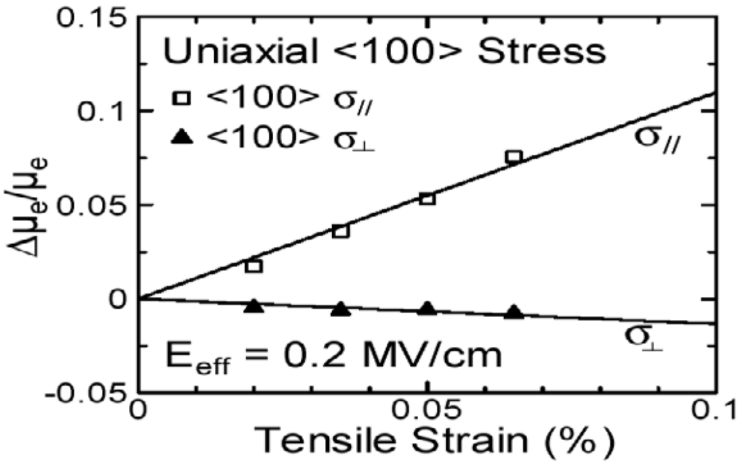



**FIGURE 4.53**
Electron mobility enhancement, $\Delta\mu_e/\mu_e$, induced by uniaxial <100> tensile stress as a function of strain. <100> $\sigma_{\|}$ enhances $\mu_e$, whereas <100> $\sigma_{\perp}$ reduces $\mu_e$. The strain is measured in the stressed direction. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

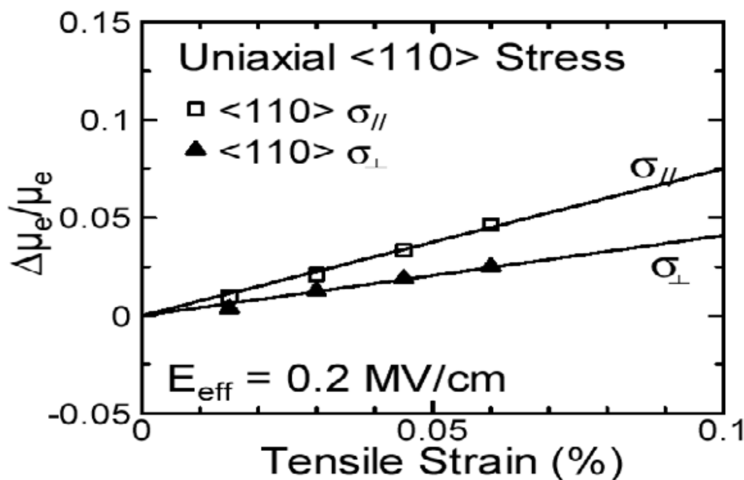



**FIGURE 4.54**
Electron mobility enhancement, $\Delta\mu_e/\mu_e$, induced by uniaxial <110> tensile stress as a function of strain. Both <110> $\sigma_{\|}$ and $\sigma_{\perp}$ enhance $\mu_e$. The strain is measured in stressed direction. After K. Uchida et al., *IEEE IEDM Tech. Dig.*, 2005(129-132). With permission.

## 4.7 Uniaxial vs. Biaxial Strain Engineering

It is important to understand the differences for biaxial and uniaxial stress and the physical mechanism responsible. The key differences between the biaxial (substrate-induced) and uniaxial (process-induced) stress have been explained and may be found in reference [55]. The superiority of using uniaxial stress over the biaxially stressed Si has been confirmed via wafer-bending experiments and also physical modeling. It has been shown that for uniaxially stressed p-MOSFETs, valence band warping creates favorable in-plane and out-of-plane low conductivity effective masses resulting in a significantly large hole mobility enhancement at low strain and at high vertical field.

Nearly all the experimental reports suggest that process-induced uniaxial stress has key advantages in both the mobility enhancement and small threshold voltage shift. It has also been shown that the uniaxial stress, unlike the biaxial stress, shows a significant hole mobility enhancement at low stress and high vertical field and also a small threshold voltage shift for n-MOSFETs [55]. To evaluate the performance enhancements via the process-induced strain, the authors have used the four-point longitudinal and transverse uniaxial and concentric ring biaxial wafer-bending experiments. For the biaxial wafer bending, material anisotropy is neglected even though silicon is orthotropic.

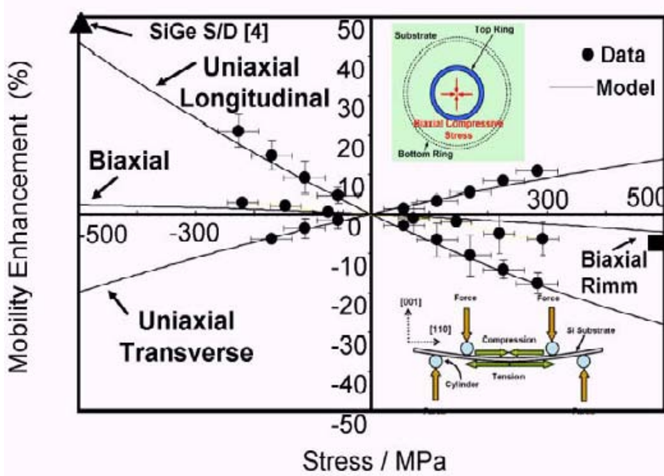



**FIGURE 4.55**
Enhanced mobility vs. stress induced via wafer bending: biaxial and longitudinal and transverse uniaxial stress. After S. E. Thompson et al., *IEEE IEDM Tech. Dig.*, 2004(221-224). With permission.

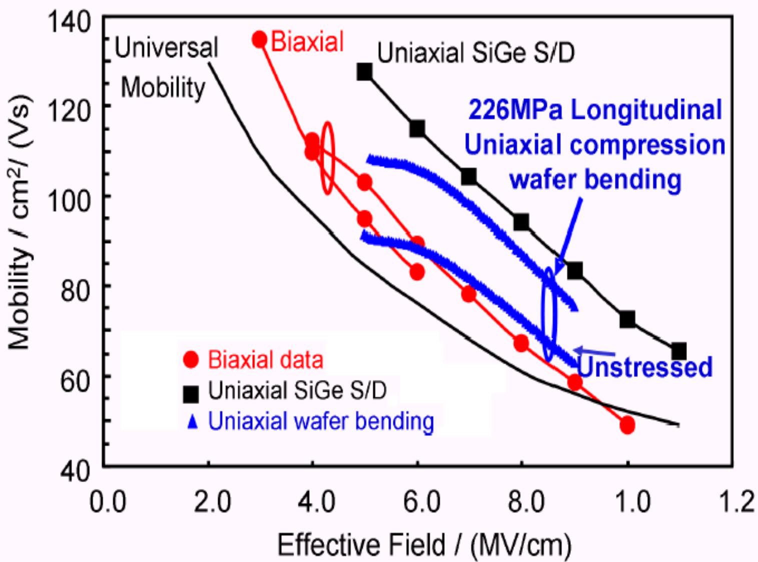



**FIGURE 4.56**
Strained-Si hole mobility enhancement vs. vertical field (wafer bending, SiGe source/drain, and biaxial substrate strain). After S. E. Thompson et al., *IEEE IEDM Tech. Dig.*, 2004(221-224). With permission.

Toward understanding the essential physics, the authors used a simpler formulation which consist of solving the 6×6 Luttinger-Kohn Hamiltonian to calculate the strain-altered band structure, analytical scattering models, and deformation potential theory to calculate the band edge shift needed to determine the MOSFET threshold voltage shift. The authors used the band parameters from reference [105].

The authors have shown that for both biaxial and uniaxial stressed Si, the effective mass unlike for the unstressed case, is approximately constant over the energy range of a few kT. The approximately constant effective mass results since strain removes the degeneracy and reduces the band-to-band coupling. Also, the valence band under uniaxial compressive stress not only causes a repopulation in k-space along <110> [16] but also significantly superior in-plane conductivity mass. In-plane effective masses in the x-y plane for transport along the <110> channel direction show ∼ 40% lighter effective mass for uniaxial stress. It is this superior in-plane mass for uniaxial stress that results in enhanced mobility at low strain. Changes in scattering also contribute to the hole mobility improvement but has been shown to be a small at uniaxial stress levels of (∼ 500 MPa) [11] and even larger (∼ 900 MPa) [4].

For the six possible types of stresses, the mobility extracted from changes in the linear drain current is shown in Fig. 4.55. Transistors from a standard 90-nm technology [3, 4, 11, 16]. were used in the wafer-bending measurements. In general, the wafer bending data support the biaxial substrate and uniaxial process-induced stress observations. As shown in Fig. 4.56, much larger hole mobility enhancement at low and moderate stress on p-MOSFETs is observed at both low and high vertical fields, which is consistent with the uniaxial process-induced stress data. Wafer-bending uniaxial tensile stress also shows a small threshold voltage shift for n-MOSFETs.

Careful analysis [101] of the 1990's biaxial and recent uniaxial strained-Si experimental data suggest the following key advantages of process-induced uniaxial strain as follows:

- Uniaxial stress provides significantly larger hole mobility enhancement at both low strain and high vertical electric field due to differences in the warping of the valence band under strain. Large mobility enhancement at low strain is important since yield loss via dislocations occurs at high strain.
- Uniaxial stress enhanced mobility provides larger drive current improvement for nanoscale short-channel devices. This is due to fact that the uniaxial stress-enhanced electron and hole mobility arises mostly from reduced conductivity effective mass (vs. reduced scattering for biaxial stress), since uniaxial shear stress provides significant valence and some conduction band warping.
- Process-induced uniaxial stress causes approximately five times smaller n-channel threshold voltage shift compared to substrate-induced biaxial tensile stress.

Process-induced strain techniques are being integrated in a modular fashion into any existing CMOS baseline process for integrated circuit manufacturing. Extra process complexity, cost, and yield should be carefully considered (see

**TABLE 4.2**
Summary of all feasible technologies for mobility enhancement. Extra processes are considered, such as selective etch or implantation to prevent performance degradation. After V. Chan et al., *IEEE CICC Conf. Dig.*, 2005(667-674). With permission.

| | nMOS | pMOS | Disadvantage and limitations |
|---|---|---|---|
| Biaxial Tensile Strain | ↑ | ↑ | Extra cost on substrate, difficulty in material, integration, and device design |
| Contact etch-stop liner (DSL) | ↑ | ↑ | Extra steps in integration, ground rule consideration |
| SMT | ↑ | o | Extra steps in integration |
| e-SiGe | o | ↑ | High process complexity difficulty in epi-growth, yield |
| Substrate Orientation (HOT) | o | ↑ | Extra Cost on hybrid substrate, extra steps e.g., Epi to prepare isolation |
| Channel Orientation (<100>) | o | ↑ | No further improvement in PMOS current |

Table 4.2). In addition, extra process steps may involve extra film deposition or etch. Extra plasma damage and low temperature process may lead to reliability issues.

## 4.8 Summary

In this chapter, electronic properties of biaxial tensilely strained-Si films have been discussed. Material properties such as bandgap, lattice constants, velocity-field characteristics, conduction and valence band discontinuities, heavy doping effect, and carrier mobilities of SiGe, SiGeC, and strained-Si layers have been presented. The essential physics explaining the physical origin of the differences in the quantum confinement for uniaxial and biaxial stress for Si and Ge is described.

Process-induced strain is achieved by adjusting the tensile stress of the nitride cap layer (via thickness variation) for n-MOSFETs and longitudinal uniaxial compressive strain by Ge content variation in the source/drain region for p-MOSFETs. The uniaxial strained-Si p-MOSFET is seen to have significant advantages over the biaxial strain, since (a) it has a much larger mobility enhancement for a given strain, and (b) the mobility enhancement is present at high vertical fields. It is expected that uniaxial stress will be the dominant type of stress adopted in logic technologies due to advantages observed for

both n- and p-MOSFETs. Since uniaxial stress hole mobility enhancement results from a mass instead of a scattering improvement, the enhanced hole mobility improvement will be present in nanoscale ballistic transport-limited MOSFETs. In process-induced stress, since many process flow parameters are changed when fabricating strained-Si MOSFETs, there is some uncertainty if the process-induced strain alone is responsible for the reported enhancements both in mobility and threshold voltage shift.

## References

[1] C. K. Maiti and G. A. Armstrong, *Applications of Silicon-Germanium Heterostructure Devices.* Institute of Physics Publishing, UK, 2001.

[2] S. Thompson, N. Anand, M. Armstrong, C. Auth, B. Arcot, M. Alavi, P. Bai, J. Bielefeld, R. Bigwood, J. Brandenburg, M. Buehler, S. Cea, V. Chikarmane, C. Choi, R. Frankovic, T. Ghani, G. Glass, W. Han, T. Hoffmann, M. Hussein, P. Jacob, A. Jain, C. Jan, S. Joshi, C. Kenyon, J. Klaus, S. Klopcic, J. Luce, Z. Ma, B. Mcintyre, K. Mistry, A. Murthy, P. Nguyen, H. Pearson, T. Sandford, R. Schweinfurth, R. Shaheed, S. Sivakumar, M. Taylor, B. Tufts, C. Wallace, P. Wang, C. Weber, and M. Bohr, "A 90 nm logic technology featuring 50 nm strained silicon channel transistors, 7 layers of Cu interconnects, low-k ILD, and 1 $\mu$m$^2$ SRAM cell," *IEEE IEDM Tech. Dig.*, pp. 61–64, 2002.

[3] V. Chan, R. Rengarajan, N. Rovedo, W. Jin, T. Hook, P. Nguyen, J. Chen, E. Nowak, X.-D. Chen, D. Lea, A. Chakravarti, V. Ku, S. Yang, A. Steegen, C. Baiocco, P. Shafer, H. Ng, S.-F. Huang, and C. Wann, "High speed 45 nm gate length CMOSFETs integrated into a 90 nm bulk technology incorporating strain engineering," *IEEE IEDM Tech. Dig.*, pp. 77–80, 2003.

[4] P. R. Chidambaram, B. A. Smith, L. H. Hall, H. Bu, S. Chakravarthi, Y. Kim, A. V. Samoilov, A. T. Kim, P. J. Jones, R. B. Irwin, M. J. Kim, A. L. P. Rotondaro, C. F. Machala, and D. T. Grider, "35% drive current improvement from recessed-SiGe drain extensions on 37 nm gate length PMOS," *IEEE VLSI Tech. Symp. Dig.*, pp. 48–49, 2004.

[5] S. Ito, H. Namba, K. Yamaguchi, T. Hirata, K. Ando, S. Koyama, S. Kuroki, N. Ikezawa, T. Suzuki, T. Saitoh, and T. Horiuchi, "Mechanical stress effect of etch-stop nitride and its impact on deep submicron transistor design," *IEEE IEDM Tech. Dig.*, pp. 247–250, 2000.

[6] A. Shimizu, K. Hachimine, N. Ohki, H. Ohta, M. Koguchi, Y. Nonaka, H. Sato, and F. Ootsuka, "Local mechanical-stress control (LMC): a new

technique for CMOS-performance enhancement," *IEEE IEDM Tech. Dig.*, pp. 433–436, 2001.

[7] C. K. Maiti, L. K. Bera, S. S. Dey, D. K. Nayak, and N. B. Chakrabarti, "Hole mobility enhancement in strained-Si p-MOSFETs under high vertical fields," *Solid-State Electron.*, vol. 41, pp. 1863–1869, 1997.

[8] K. Rim, J. Chu, H. Chen, K. A. Jenkins, T. Kanarsky, K. Lee, A. Mocuta, H. Zhu, R. Roy, J. Newbury, J. Ott, K. Petrarca, P. Mooney, D. Lacey, S. Koester, K. Chan, D. Boyd, M. Leong, and H.-S. Wong, "Characteristics and device design of sub-100 nm strained-Si n- and p-MOSFETs," *IEEE VLSI Tech. Symp. Dig.*, pp. 98–99, 2002.

[9] J. L. Hoyt, H. M. Nayfeh, S. Eguchi, I. Aberg, G. Xia, T. Drake, E. A. Fitzgerald, and D. A. Antoniadis, "Strained silicon MOSFET technology," *IEEE IEDM Tech. Dig.*, pp. 23–26, 2002.

[10] M. L. Lee and E. A. Fitzgerald, "Hole mobility enhancements in nanometer-scale strained-silicon heterostructures grown on Ge-rich relaxed $Si_{1-x}Ge_x$," *J. Appl. Phys.*, vol. 94, pp. 2590–2596, 2003.

[11] S. E. Thompson, M. Armstrong, C. Auth, S. Cea, R. Chau, G. Glass, T. Hoffman, J. Klaus, Z. Ma, B. Mcintyre, A. Murthy, B. Obradovic, L. Shifren, S. Sivakumar, S. Tyagi, T. Ghani, K. Mistry, M. Bohr, and Y. El-Mansy, "A logic nanotechnology featuring strained-silicon," *IEEE Electron Dev. Lett.*, vol. 25, pp. 191–193, 2004.

[12] K. Natori, "Ballistic metal-oxide-semiconductor field effect transistor," *J. Appl. Phys.*, vol. 76, pp. 4879–4890, 1994.

[13] M. Lundstrom and Z. Ren, "Essential physics of carrier transport in nanoscale MOSFETs," *IEEE Trans. Electron Dev.*, vol. 49, pp. 133–141, 2002.

[14] S. Takagi, "Re-examination of subband structure engineering in ultrashort channel MOSFETs under ballistic carrier transport," *IEEE VLSI Tech. Symp. Dig.*, pp. 115–117, 2003.

[15] A. Rahman and M. S. Lundstrom, "A compact scattering model for the nanoscale double-gate MOSFET," *IEEE Trans. Electron Dev.*, vol. 49, pp. 481–489, 2002.

[16] M. D. Giles, M. Armstrong, C. Auth, S. M. Cea, T. Ghani, T. Hoffmann, R. Kotlyar, P. Matagne, K. Mistry, R. Nagisetty, B. Obradovic, R. Shaheed, L. Shifren, M. Stettler, S. Tyagi, X. Wang, C. Weber, and K. Zawadzki, "Understanding stress enhanced performance in Intel 90 nm CMOS technology," *IEEE VLSI Tech. Symp. Dig.*, pp. 118–119, 2004.

[17] C. K. Maiti, N. B. Chakrabarti, and S. K. Ray, *Strained Silicon Heterostructures: Materials and Devices.* The IEE, UK, 2001.

[18] C. K. Maiti, L. K. Bera, and S. Chattopadhyay, "Strained-Si heterostructure field effect transistors," *Semicond. Sci. Technol.*, vol. 13, pp. 1225–1246, 1998.

[19] J. Welser, J. L. Hoyt, and J. F. Gibbons, "Electron mobility enhancement in strained-Si N-type metal-oxide-semiconductor field-effect transistors," *IEEE Electron. Dev. Lett.*, vol. 15, pp. 100–102, 1994.

[20] J. Welser, J. L. Hoyt, S. Takagi, and J. F. Gibbons, "Strain Dependence of the performance enhancement in strained-Si n-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 373–376, 1994.

[21] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Transconductance enhancement in deep submicron strained-Si n-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 707–710, 1998.

[22] T. Mizuno, N. Sugiyama, T. Tezuka, T. Numata, and S. Takagi, "High performance CMOS operation of strained-SOI MOSFETs using thin film SiGe-on-insulator," *IEEE VLSI Tech. Symp. Dig.*, pp. 106–109, 2002.

[23] K. Rim, S. Koester, M. Hargrove, J. Chu, P. M. Mooney, J. Ott, T. Kanarsky, P. Ronsheim, M. Ieong, A. Grill, and H.-S. P. Wong, "Strained Si nMOSFETs for high performance CMOS technology," *VLSI Technology*, pp. 59–60, 2001.

[24] M. S. Lundstrom, "On the mobility versus drain current relation for a nanoscale MOSFET," *IEEE Electron Dev. Lett.*, vol. 22, pp. 293–295, 2001.

[25] A. Lochtefeld and D. A. Antoniadis, "Investigating the relationship between electron mobility and velocity in deeply scaled nMOS via mechanical stress," *IEEE Electron Dev. Lett.*, vol. 22, pp. 591–593, 2001.

[26] J. L. Egley and D. Chidambarrao, "Strain effects on device characteristics: implementation in drift-diffusion simulators," *Solid-State Electron.*, vol. 36, pp. 1653–1664, 1993.

[27] M. V. Fischetti and S. E. Laux, "Band structure, deformation potentials, and carrier mobility in strained-Si, Ge, and SiGe alloys," *J. Appl. Phys.*, vol. 88, pp. 2234–2252, 1996.

[28] M. V. Fischetti, Z. Ren, P. M. Solomon, M. Yang, and K. Rim, "Six-band k·p calculation of the hole mobility in silicon inversion layers: dependence on surface orientation, strain, and silicon thickness," *J. Appl. Phys.*, vol. 94, pp. 1079–1095, 2003.

[29] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, "Hole mobility enhancements and alloy scattering-limited mobility in tensile strained-Si/SiGe surface channel metal-oxide-

semiconductor field-effect transistors," *J. Appl. Phys.*, vol. 92, pp. 3745–3751, 2002.

[30] J. R. Chelikowsky and M. L. Cohen, "Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors," *Phys. Rev. B*, vol. 14, pp. 556–582, 1976.

[31] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev. B*, vol. 136, pp. 864–871, 1964.

[32] S. Froyen, D. M. Wood, and A. Zunger, "Structural and electronic properties of epitaxial thin layer superlattices," *Phys. Rev. B*, vol. 37, pp. 6893–6907, 1988.

[33] R. People and J. C. Bean, "Band alignments of coherently strained $Ge_xSi_{1-x}$/Si heterostructures on <001> $Ge_ySi_{1-y}$ substrates," *Appl. Phys. Lett.*, vol. 48, pp. 538–540, 1986.

[34] R. Braunstein, A. R. Moore, and F. Herman, "Intrinsic optical absorption in germanium-silicon alloys," *Phys. Rev.*, vol. 109, pp. 695–710, 1958.

[35] J. Weber and M. I. Alonso, "Near-band-gap photoluminescence of SiGe alloys," *Phys. Rev. B*, vol. 40, pp. 5683–5693, 1989.

[36] R. People, "Indirect band gap of coherently strained $Ge_xSi_{1-x}$ bulk alloys on <100> silicon substrates," *Phys. Rev.*, vol. 32, pp. 1405–1408, 1985.

[37] R. People, "Indirect band gap and band alignment for coherently strained $Si_{1-x}Ge_x$ bulk alloys on germanium (001) substrates," *Phys. Rev. B*, vol. 34, pp. 2508–2510, 1986.

[38] C. G. V. de Walle and R. M. Martin, "Theoretical calculations of heterojunction discontinuities in the Si/Ge system," *Phys. Rev. B*, vol. 34, pp. 5621–5634, 1986.

[39] J. C. Bean, "Silicon-based semiconductor heterostructures: column IV bandgap engineering," *Proc. IEEE*, vol. 80, pp. 571–587, 1992.

[40] A. Rahman, G. Klimeck, T. B. Boykin, and M. Lundstrom, "Bandstructure effects in ballistic nanoscale MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 139–142, 2004.

[41] T. Manku and A. Nathan, "Electron drift mobility model for devices based on unstrained and coherently strained-$Si_{1-x}Ge_x$ grown on <001> silicon substrate," *IEEE Trans. Electron Dev.*, vol. 39, pp. 2082–2089, 1992.

[42] C. Herring and E. Vogt, "Transport and deformation-potential theory for many-valley semiconductor with anisotropic scattering," *Phys. Rev.*, vol. 101, pp. 944–961, 1956.

[43] R. People, "Physics and applications of $\mathrm{Ge}_x\mathrm{Si}_{1-x}$ strained-layer heterostructures," *IEEE J. Quantum Elec.*, vol. 22, pp. 1696–1710, 1986.

[44] R. People and J. C. Bean, "Band alignments of coherently strained $\mathrm{Ge}_x\mathrm{Si}_{1-x}$/Si heterostructures on (001) $\mathrm{Ge}_y\mathrm{Si}_{1-y}$ substrate," *Appl. Phys. Lett.*, vol. 48, pp. 1338–1340, 1986.

[45] U. Konig and H. Daembkes, "SiGe HBTs and HFETs," *Solid-State Electron.*, vol. 38, pp. 1595–1602, 1995.

[46] D. K. Nayak and S. K. Chun, "Low-field hole mobility of strained-Si on (100) $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ substrate," *Appl. Phys. Lett.*, vol. 64, pp. 2514–2516, 1994.

[47] F. Schaffler, "High-mobility Si and Ge structures," *Semicond. Sci. Technol.*, vol. 12, pp. 1515–1549, 1997.

[48] S. C. Jain and D. J. Roulston, "A simple expression for band gap narrowing (BGN) in heavily doped Si, Ge, GaAs and $\mathrm{Ge}_x\mathrm{Si}_{1-x}$ strained layers," *Solid-State Electron.*, vol. 34, pp. 453–465, 1991.

[49] J. S. Yuan, *SiGe, GaAs, and InP Heterojunction Bipolar Transistors.* John Wiley & Sons, New York, 1999.

[50] Z. Matutinovic-Krstelj, V. Venkataraman, E. J. Prinz, J. C. Sturm, and C. W. Magee, "Base resistance and effective bandgap reduction in n-p-n Si/$\mathrm{Si}_{1-x}\mathrm{Ge}_x$/Si HBTs with heavy base doping," *IEEE Trans. Electron Dev.*, vol. 43, pp. 457–466, 1996.

[51] B. K. Ridley, *Quantum Processes in Semiconductors.* Oxford University Press, New York, 3rd ed., 1993.

[52] J. M. Hinckley and J. Singh, "Hole transport theory in pseudomorphic $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ alloys grown on Si(001) substrates," *Phys. Rev. B*, vol. 41, pp. 2912–2926, 1990.

[53] Z.-Y. Cheng, M. T. Currie, C. W. Leitz, G. Tarashci, E. A. Fitzgerald, J. L. Hoyt, and D. A. Antoniadis, "Electron mobility enhancement in strained-Si n-MOSFETs fabricated on SiGe-on-insulator (SGOI) substrates," *IEEE Electron Dev. Lett.*, vol. 22, pp. 321–323, 2001.

[54] V. Venkataraman, C. W. Liu, and J. C. Sturm, "Alloy scattering limited transport of two-dimensional carriers in strained-$\mathrm{Si}_{1-x}\mathrm{Ge}_x$ quantum wells," *Appl. Phys. Lett.*, vol. 63, pp. 2795–2797, 1993.

[55] S. E. Thompson, G. Sun, K. Wu, J. Kim, and T. Nishida, "Key differences for process-induced uniaxial vs. substrate-induced biaxial stressed Si and Ge channel MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 221–224, 2004.

[56] R. W. Keyes, "High-mobility FET in strained silicon," *IEEE Trans. Electron Dev.*, vol. ED-33, p. 863, 1986.

[57] Th. Vogelsang and K. R. Hofmann, "Electron transport in strained-Si layers on $Si_{1-x}Ge_x$ substrates," *Appl. Phys. Lett.*, vol. 63, pp. 186–188, 1993.

[58] J. B. Roldan, F. Gamiz, J. A. Lopez-Villanueva, and J. E. Carceller, "A Monte Carlo study on the electron-transport properties of high performance strained-Si on relaxed $Si_{1-x}Ge_x$ channel MOSFETs," *J. Appl. Phys.*, vol. 80, pp. 5121–5128, 1996.

[59] G. Abstreiter, T. Wolf, H. Jorke, and H. J. Herzog, "Strain-induced two-dimensional electron gas in selectively doped $Si/Si_xGe_{1-x}$ superlattice," *Phys. Rev. Lett.*, vol. 54, pp. 2441–2445, 1985.

[60] N. Mohta and S. E. Thompson, "Mobility enhancement - the next vector to extend Moore's law," *IEEE Circuits and Dev. Mag.*, vol. September/October, pp. 18–23, 2005.

[61] Th. Vogelsang and K. R. Hofmann, "Electron mobilities and high-field drift velocities in strained silicon on silicon-germanium substrates," *IEEE Dev. Res. Conf. Dig.*, pp. 34–35, 1992.

[62] F. Schaffler, D. Tobben, H. J. Herzog, G. Abstreiter, and B. Hollander, "High-electron-mobility Si/SiGe heterostructures: influence of the relaxed SiGe buffer layer," *Semicond. Sci. Technol.*, vol. 7, pp. 260–266, 1992.

[63] Y. J. Mii, Y. H. Xie, E. A. Fitzgerald, D. Monroe, F. A. Thiel, B. Weir, and L. C. Feldman, "Extremely high electron mobility in $Si/Ge_xSi_{1-x}$ structures grown by molecular beam epitaxy," *Appl. Phys. Lett.*, vol. 59, pp. 1611–1613, 1991.

[64] R. Oberhuber, G. Zandler, and P. Vogl, "Subband structure and mobility of two-dimensional holes in strained-Si/SiGe MOSFETs," *Phys. Rev. B*, vol. 58, pp. 9941–9948, 1998.

[65] H. Nakatsuji, Y. Kamakura, and K. Taniguchi, "A study of subband structure and transport of two-dimensional holes in strained-Si p-MOSFETs using full-band modeling," *IEEE IEDM Tech. Dig.*, pp. 727–730, 2002.

[66] M. V. Fischetti and S. Laux, "Monte Carlo study of electron transport in silicon inversion layers," *Phys. Rev. B*, vol. 48, pp. 2244–2274, 1993.

[67] M. M. Rieger and P. Vogl, "Electronic-band parameters in strained $Si_{1-x}Ge_x$ alloys on $Si_{1-y}Ge_y$ substrates," *Phys. Rev. B*, vol. 48, pp. 14276–14287, 1993.

[68] A. Pirovano, A. L. Lacaita, G. Zandler, and R. Oberhuber, "Explaining the dependences of electron and hole mobilities in Si MOSFETs inversion layer," *IEEE IEDM Tech. Dig.*, pp. 527–530, 1999.

[69] A. L. Fetter, "Electrodynamics and thermodynanics of a classical electron surface layer," *Phys. Rev. B*, vol. 10, pp. 3739–3745, 1974.

[70] S. Takagi, A. Toriumi, M. Iwase, and H. Tango,, "On the universality of inversion layer mobility in Si MOSFETs: Part I - Effects of substrate impurity concentration," *IEEE Trans. Electron Dev.*, vol. 41, pp. 2357–2362, 1994.

[71] K. Rim, R. Anderson, D. Boyd, F. Cardone, K. Chan, H. Chen, S. Christansen, J. Chu, K. Jenkins, T. Kanarsky, S. Koester, B. H. Lee, K. Lee, V. Mazzeo, A. Mocuta, D. Mocuta, P. M. Mooney, P. Oldiges, J. Ott, P. Ronsheim, R. Roy, A. Steegen, M. Yang, H. Zhu, M. Ieong, and H.-S. P. Wong, "Strained Si CMOS (SS CMOS) technology: opportunities and challenges," *Solid-State Electron.*, vol. 47, pp. 1133–1139, 2003.

[72] S. Takagi, J. L. Hoyt, J. J. Welser, and J. F. Gibbons, "Comparative study of phonon-limited mobility to two-dimensional electrons in strained and unstrained Si metal-oxide-semiconductor field-effect transistors," *J. Appl. Phys.*, vol. 80, pp. 1567–1577, 1996.

[73] M. V. Fischetti, F. Gamiz, and W. E. Haensch, "On the enhanced electron mobility in strained-silicon inversion layers," *J. Appl. Phys.*, vol. 92, pp. 7320–7324, 2002.

[74] T. Mizuno, N. Sugiyama, T. Tezuka, Y. Moriyama, S. Nakaharai, T. Maeda, and S. Takagi, "Physical mechanism for high hole mobility in (110)-surface strained- and unstrained-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 809–812, 2003.

[75] T. Mizuno, N. Sugiyama, T. Tezuka, T. Numata, and S. Takagi, "Thin-film strained-SOI CMOS devices - physical mechanisms for reduction of carrier mobility," *IEEE Trans. Electron Dev.*, vol. 51, pp. 1114–1121, 2004.

[76] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Fabrication and analysis of deep submicron strained-Si N-MOSFETs," *IEEE Trans. Electron Dev.*, vol. 47, pp. 1406–1415, 2000.

[77] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Fabrication and analysis of deep submicron strained-Si n-MOSFETs," *IEEE Trans. Electron Dev.*, vol. 47, pp. 281–286, 2000.

[78] G. G. Shahidi, D. A. Antoniadis, and H. I. Smith, "Electron velocity overshoot at room and liquid nitrogen temperatures in silicon inversion layers," *IEEE Electron Dev. Lett.*, vol. 9, pp. 94–96, 1988.

[79] H. M. Nayfeh, C. W. Leitz, A. J. Pitera, E. A. Fitzgerald, J. L. Hoyt, and D. A. Antoniadis, "Influence of high channel doping on the inversion

layer electron mobility in strained silicon n-MOSFETs," *IEEE Electron Dev. Lett.*, vol. 24, pp. 248–250, 2003.

[80] F. Gamiz, J. B. Roldan, J. A. L. Villanueva, and P. Cartujo, "Coulomb scattering in strained-silicon inversion layers on $Si_{1-x}Ge_x$ substrates," *Appl. Phys. Lett.*, vol. 69, pp. 797–799, 1996.

[81] M.-S. Liang, J. Y. Choi, P.-K. Ko, and C. Hu, "Inversion-layer capacitance and mobility of very thin gate-oxide MOSFETs," *IEEE Trans. Electron Dev.*, vol. 33, pp. 409–412, 1986.

[82] J. W. Matthews and A. E. Blakeslee, "Defects in epitaxial multilayers - I. misfit dislocations," *J. Crystal Growth*, vol. 27, pp. 118–125, 1974.

[83] J. G. Fiorenza, G. Braithwaite, C. W. Leitz, M. T. Currie, J. Yap, F. Singaporewala, V. K. Yang, T. A. Langdo, J. Carlin, M. Somerville, A. Lochtefeld, H. Badawi, and M. T. Bulsara, "Film thickness constraints for manufacturable strained silicon CMOS," *Semicond. Sci. Technol.*, vol. 19, pp. L4–L8, 2004.

[84] T. Mizuno, N. Sugiyama, T. Tezuka, T. Numata, and S. Takagi, "High-performance strained-SOI CMOS devices using thin film SiGe-on-insulator technology," *IEEE Trans. Electron Dev.*, vol. 50, pp. 988–994, 2003.

[85] S. F. Nelson, K. Ismail, J. O. Chu, and B. S. Meyerson, "Room-temperature electron mobility in strained-Si/SiGe heterostructures," *Appl. Phys. Lett.*, vol. 63, pp. 367–369, 1993.

[86] F. Gamiz, J. B. Roldan, and A. Godoy, "Strained-Si/strained-Si-on-insulator inversion layers: the role of strained-Si layer thickness on electron mobility," *Appl. Phys. Lett.*, vol. 80, pp. 4160–4162, 2002.

[87] S. B. Samavedam, W. J. Taylor, J. M. Grant, J. A. Smith, P. J. Tobin, A. Dip, A. M. Phillips, and R. Liu, "Relaxation of strained-Si layers grown on SiGe buffers," *J. Vac. Sci. Technol. B*, pp. 1424–1429, 1999.

[88] N. Sugii, "Thermal stability of the strained-Si/$Si_{0.7}Ge_{0.3}$ heterostructure," *J. Appl. Phys.*, vol. 89, pp. 6459–6463, 2001.

[89] S. J. Koester, K. Rim, J. O. Chu, P. M. Mooney, J. A. Ott, and M. A. Hargrove, "Effect of thermal processing on strain relaxation and interdiffusion in Si/SiGe heterostructures studied using Raman spectroscopy," *Appl. Phys. Lett.*, vol. 79, pp. 2148–2150, 2001.

[90] P. Kringhoj, A. N. Larsen, and S. Y. Shirayev, "Diffusion of Sb in strained and relaxed Si and SiGe," *Phys. Rev. Lett.*, vol. 76, pp. 3372–3375, 1996.

[91] N. R. Zangenberg, J. Fage-Pedersen, J. L. Hansen, and A. N. Larsen, "Boron and phosphorus diffusion in strained and relaxed Si and SiGe," *J. Appl. Phys.*, vol. 94, pp. 3883–3890, 2003.

[92] P. Laitinen, I. Riihimaki, and J. Raisanen, "Arsenic diffusion in relaxed $Si_{1-x}Ge_x$," *Phys. Rev. B*, vol. 68, pp. 155209–155214, 2003.

[93] S. Eguchi, C. N. Chleirigh, O. O. Olubuyide, and J. L. Hoyt, "Germanium-concentration dependence of arsenic diffusion in silicon germanium alloys," *Appl. Phys. Lett.*, pp. 368–370, 2004.

[94] MIT, "MIT Annual Report," 2002.

[95] K. Suzuki, H. Hasegawa, and Y. Kanda, "Origin of the linear and non-linear piezoresistance effects in p-type silicon," *Jpn. J. Appl. Phys.*, pp. 871–872, 1984.

[96] C. S. Smith, "Piezoresistance effect in germanium and silicon," *Phys. Rev.*, vol. 94, pp. 42–49, 1954.

[97] G. Dorda, "Piezoresistance in quantized conduction bands in silicon inversion layers," *J. Appl. Phys.*, vol. 42, pp. 2053–2060, 1971.

[98] A. Steegen, M. Stucchi, A. Lauwers, and K. Maex, "Silicide induced pattern density and orientation dependent transconductance in MOS transistor," *IEEE IEDM Tech. Dig.*, pp. 497–500, 1999.

[99] G. E. Pikus and G. L. Bir, *Symmetry and Strain-Induced Effects in Semiconductors.* Wiley, New York, 1974.

[100] R. E. Belford, Z. Wei, J. Potashnik, L. Qingmin, and A. Seabaugh, "Performance-augmented CMOS using back-end uniaxial strain," *IEEE DRC Proc.*, pp. 41–42, 2002.

[101] S. E. Thompson, G. Sun, Y. S. Choi, and T. Nishida, "Uniaxial-process-induced strained-Si: extending the CMOS roadmap," *IEEE Trans. Electron Dev.*, vol. 53, pp. 1010–1020, 2006.

[102] E. Wang, P. Matagne, L. Shifren, B. Obradovic, R. Kotlyar, S. Cea, J. He, Z. Ma, R. Nagisetty, S. Tyagi, M. Stettler, and M. D. Giles, "Quantum mechanical calculation of hole mobility in silicon inversion layers under arbitrary stress," *IEEE IEDM Tech. Dig.*, pp. 147–150, 2004.

[103] B. Obradovic, P. Matagne, L. Shifren, X. Wang, M. Stettler, J. He, and M. D. Giles, "A physically-based analytic model for stress-induced hole mobility enhancement," *IWCE-10 Tech. Dig.*, pp. 26–27, 2004.

[104] K. Uchida, T. Krishnamohan, K. C. Saraswat, and Y. Nishi, "Physical mechanisms of electron mobility enhancement in uniaxial stressed MOSFETs and impact of uniaxial stress engineering in ballistic regime," *IEEE IEDM Tech. Dig.*, pp. 129–132, 2005.

[105] P. Lawaetz, "Valence-band parameters in cubic semiconductors," *Phys. Rev. B*, vol. 4, pp. 3460–3467, 1971.

# 5

## *Gate Dielectrics on Engineered Substrates*

The use of high-quality strained-Si layers grown on relaxed-SiGe substrates or on an insulator holds a great promise to enhance conventional Si complementary metal oxide semiconductor (Si CMOS) technology. In order to fabricate MOSFETs on a Si substrate, conventionally a very thin layer of insulator is thermally grown or deposited on the Si substrate followed by a polysilicon gate electrode definition. With the advent of strained-$Si_{1-x}Ge_x$-based devices, it is important to consider some of the processing issues that separate these alloy layers from conventional Si substrate and conventional Si processing technology. One area of concern in the metal oxide semiconductor-based $Si_{1-x}Ge_x$ technologies is the formation of high-quality surface passivation with low density of interface states.

Attempts to thermally passivate alloys of $Si_{1-x}Ge_x$, using conventional thermal oxidation, result in the selective oxidation of the Si in the alloy and the resulting formation of a layer of Ge at the oxide/alloy interface, which greatly increases the interface trap density [1]. In addition, if the strained layer is mechanically metastable, then the high temperatures needed for conventional oxidation may result in the introduction of "strain-relieving defects" at the strained-layer/substrate interface. The resulting Ge buildup at the oxide/alloy interface and strain relaxation at the alloy/substrate interface render the heterostructure useless for electronic device applications. With the realization of heterojunction bipolar $Si_{1-x}Ge_x$ devices and the potential for optoelectronic devices based on these material systems comes the need for a low-temperature technique that produces high-quality surface passivation.

The thickness of the gate oxide determines the threshold voltage that must be applied at the gate of a MOSFET to turn on the MOSFET device. As the operational speed of electrical systems has increased, it has become necessary to have MOSFET devices with different gate oxide thicknesses on the same chip. It is also important to have the scalability of the gate dielectric layers on Si heterolayers.

The purpose of this chapter is to discuss the developments on the gate dielectric formation on strained-Si/SiGe heterolayers. The processing issues of gate dielectric formation on strained-Si films are critically examined and the thermal oxidation of strained-Si layers is discussed. Low thermal budget processing, such as rapid thermal oxidation (RTO), and low temperature microwave plasma oxidation and nitridation of strained-Si layers are discussed.

A brief review is given for conventional thermal oxidation of strained-Si,

and readers are referred to the original publications for more details. We give a more in-depth discussion on the relatively new results: necessity for low thermal budget processing for the fabrication of strained-Si channel MOSFETs, including rapid thermal oxidation and nitridation, and microwave plasma oxidation. Recent progress made in terms of the integration issues and the prospects of deposition of high-k gate dielectrics on strained-Si/SiGe heterolayers will be given more emphasis for possible applications in high-performance hetero-FETs. Microwave plasma deposition of various high-k gate dielectrics, such as $ZrO_2$, $Ta_2O_5$, and $TiO_2$ on strained-Si, their electrical properties, and the current conduction mechanisms are also discussed. Results of studies on the electrical properties and conduction mechanisms in several high-k gate dielectrics, such as $ZrO_2$, $Ta_2O_5$, and $TiO_2$ deposited on strained-Si are also presented.

## 5.1 Strained-Si MOSFET Structures

Surface-channel strained-Si MOSFETs on relaxed-$Si_{1-x}Ge_x$ virtual substrates are attractive for extending the Si CMOS platform. The performance of a surface-channel MOSFET depends on the Si/$SiO_2$ interface quality. As discussed earlier, the oxidation of ultrathin (100-200 Å) strained-Si layer is a key processing issue as the gate oxide thickness and strained-Si channel thickness (remaining strained-Si layer after oxidation) are determined by this gate oxidation step. High-temperature ($> 700^oC$) processing is not desirable for strained-Si/SiGe structures to prevent Ge diffusion, strain relaxation, and dislocation propagation from the buffer layers. Many of the reported hetero-FET structures use a MODFET configuration with a Schottky gate to avoid the high-temperature gate oxide growth process step [2, 3, 4]. However, several workers have used conventional thermal (both dry and wet) oxidation at a relatively low temperature (750-850$^o$C) for the fabrication of strained-Si MOSFETs [5, 6, 7, 8, 9].

First, let us consider the surface-channel strained-Si device (see Fig. 5.7). Since Si is being oxidized, the interface state density at the resulting $SiO_2$/Si interface is low, and an electrically high quality interface is expected. However, all oxidation and cleaning processes during the device fabrication will consume Si. In the case of surface channel, the top strained-Si layer is generally less than 200 Å thick, and thus too much Si consumption during cleaning and oxidation steps will etch out the high-mobility strained-Si channel.

One possible solution could be to increase the top strained-Si layer thickness to achieve the necessary thickness for the channel after processing. However, the critical thickness of the strained-Si layer, which depends on the Ge content of the relaxed-SiGe layer below, giving rise to the strain, limits the strained-Si

layer thickness. The critical thickness is a function of the lattice mismatch between the epitaxial Si film and substrate, as well as the materials properties of both the epitaxial layer and the substrate. If the critical layer thickness is exceeded, the Si layer will begin to relax, introducing misfit dislocations at the strained-Si/SiGe interface. The strain in the channel Si layer will be partially or completely relieved, decreasing the carrier mobility via scattering. Dislocations can also affect device yield, reliability, and performance.

## 5.2 Thermal Oxidation of Strained-Si

Device-quality oxides usually result from thermal oxidation of Si. Dry oxygen or steam is used at high temperature (typically 900-1000$^o$C) in the furnace, and the Si reacts with the $O_2$ to form an $SiO_2$ on the surface of the wafer. During oxidation, Si from the wafer is consumed by the growing oxide, resulting in a loss of approximately 0.44$t_{ox}$ of Si for an oxide thickness ($t_{ox}$). To investigate possible relaxation of strained-Si during thermal oxidation, Welser et al. [6] processed three strained-Si samples (grown using LRPCVD) with graded $Si_{1-x}Ge_x$ buffer layers (x = 0.30), constant Ge buffer of 1 $\mu$m thick, and varying layer thickness (160, 445, and 1020 Å) of strained-Si grown on p-substrates. The strained-Si layers were grown at 700$^o$C and the buffer layers were over 95% relaxed. The samples were doped 7×10$^{16}$ cm$^{-3}$ p-type by flowing 0.5 sccm of diborane during the growth. The samples were oxidized (both dry and wet) at temperatures ranging from 750 to 900$^o$C in a standard oxidation furnace to yield approximately 130 Å of oxide.

By measuring the wavenumber shift from Raman spectra for the strained-Si before and after oxidation Si-Si line on all of the samples, the strain before and after oxidation was found. From the Raman scan of samples oxidized at 850$^o$C, it was observed that Si-Si peak intensity from the strained-Si layer is diminished due to Si consumption during oxidation, but the unaltered peak positions of the Si-Si from strained-Si and buffer $Si_{0.7}Ge_{0.3}$ layer indicated that strain has not been reduced and relaxation remain unchanged in strained-Si and buffer $Si_{0.7}Ge_{0.3}$ layer, respectively. It was shown that for all oxidation temperatures, the thinnest strained-Si layer (160 Å) appears to remain fully strained. The 445 Å strained-Si layer is fully strained as grown, but shows some evidence of partial strain relaxation during oxidation. Finally, the thickest strained-Si layer (1020 Å) is partially relaxed even as grown, and continues to relax more as a function of oxidation temperature.

Since the oxide thickness is relatively thin ($\sim$ 130 Å), oxide growth rate is expected to be linear, and precisely that has been reported by Welser et al. [6]. Fig. 5.1 shows a plot of linear rate constants (B/A) for all three strained-Si samples and a control-Si sample as a function of temperature. The values

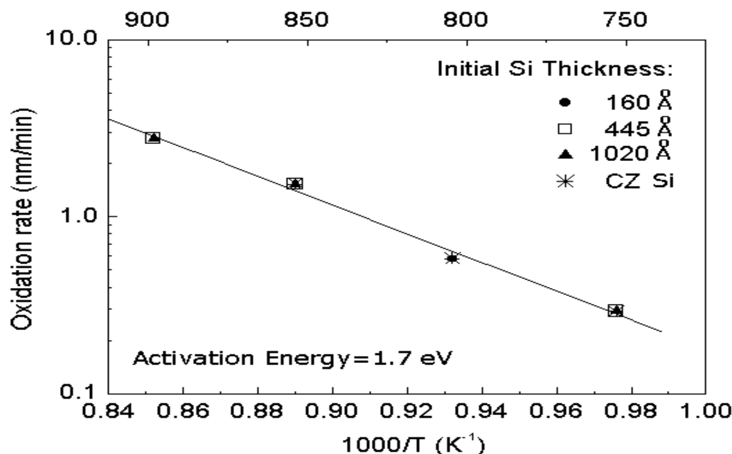



**FIGURE 5.1**
Oxide growth rate of strained-Si for wet oxidation. After J. J. Welser, *The application of strained-silicon/relaxed-silicon germanium heterostructures to metal oxide semiconductor field-effect transistors*, Ph.D. thesis, Stanford University, 1994.

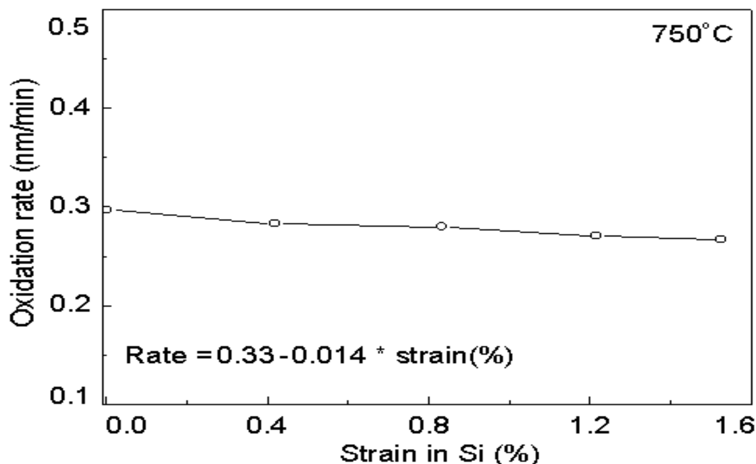



**FIGURE 5.2**
Oxidation rate vs. strain in the Si. Although the dependence is small, the oxide growth rate appears to decrease with increasing strain for wet oxidation. After J. J. Welser, *The application of strained-silicon/relaxed-silicon germanium heterostructures to metal oxide semiconductor field-effect transistors*, Ph.D. thesis, Stanford University, 1994.

extracted for all samples at all temperatures are very similar, well within the expected accuracy of the measurement. In addition, the activation energy for B/A was found to be on the order of 1.7, which is comparable to the value of 1.96 for bulk-Si.

Although there is no measured difference between the average growth rate of oxide on strained-Si as compared to bulk-Si, the oxidation rate seems to decrease slightly as strain in the Si layer increases which is shown in Fig. 5.2. All the strained-Si layers are thin and the strain in these layers varies due to the different Ge contents of the relaxed-$Si_{1-x}Ge_x$ buffer layers and the oxidation temperature is low (750$^o$C), so a very little relaxation or dislocation formation is expected during the process. Thermal (dry) oxidation of strained-Si (300 Å thick, on a fully relaxed step-graded SiGe buffer grown using GSMBE on p-Si at 700$^o$C and a quantum hole confinement have been reported by Bera et al. [10]. The authors demonstrated the hole confinement in accumulation at the $SiO_2$/strained-Si/SiGe heterointerface and reported a trap density at the middle of the bandgap of $2\times10^{12}$ cm$^{-2}$eV$^{-1}$. The authors also extracted the unconsumed strained-Si layer thickness from the capacitance-voltage characteristics [11].

### 5.2.1 Ge Diffusion

The thermal stability of the strained-Si/SiGe heterostructure is very important to enable the semiconductor industry for production of heterostructure devices and circuits. There have been several reports on the thermal stability and most of the work on Ge diffusion and strain relaxation concern SiGe/Si heterostructures and multiple quantum wells [12]. Ge atoms diffuse out through the strained-Si layer during heat treatment, which results in strain relaxation and surface roughening. The activation energy of Ge diffusion in strained-Si is 3.3 eV, which is lower than the value in unstrained-Si (4.7-5.3 eV).

The Ge depth profiles for samples treated at different temperatures are shown in Fig. 5.3(a). The strained-Si/$Si_{0.7}Ge_{0.3}$ interface was clearly observed at the depth of 25 nm for the sample with no thermal treatment. As the heat-treatment temperature increased, the Ge concentration in the strained-Si layer increased due to Ge diffusion from the $Si_{0.7}Ge_{0.3}$ layer. The Ge concentration near the surface of the sample treated at 1000$^o$C for 1 h was 1%. The concentration in the sample treated at 1100$^o$C for 1 h was almost uniform near the surface within 5-30 nm and gradually increased toward the value in the middle of the SiGe layer (30%). This suggests that a certain amount of Ge evaporated from the surface. The Ge concentration within 0-5 nm of the surface of samples treated at 950-1100$^o$C decreased because $SiO_2$ was formed during the heat treatment due to impurity oxygen gas in the furnace. It is well known that $SiO_2$ forms at the surface during thermal oxidation of SiGe. An increased oxygen concentration accompanied by a decreased Ge concentration near the surface was confirmed through Auger electron spectroscopy.

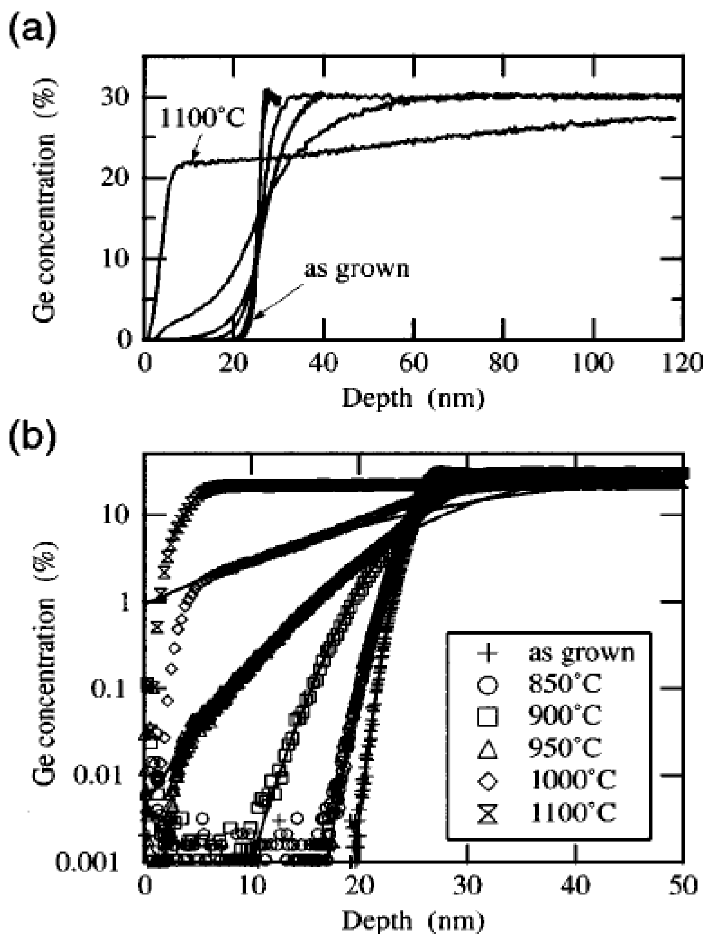



**FIGURE 5.3**
(a) Ge depth profiles (linear scale) measured by SIMS for an as-grown sample and samples treated at 850, 900, 950, 1000, or 1100$^o$C for 1 h. (b) Ge depth profiles (log scale) same as in (a) and fitting curves by Fick's law. After N. Sugii, *J. Appl. Phys.*, Vol. 89, 2001(6459-6463). With permission.

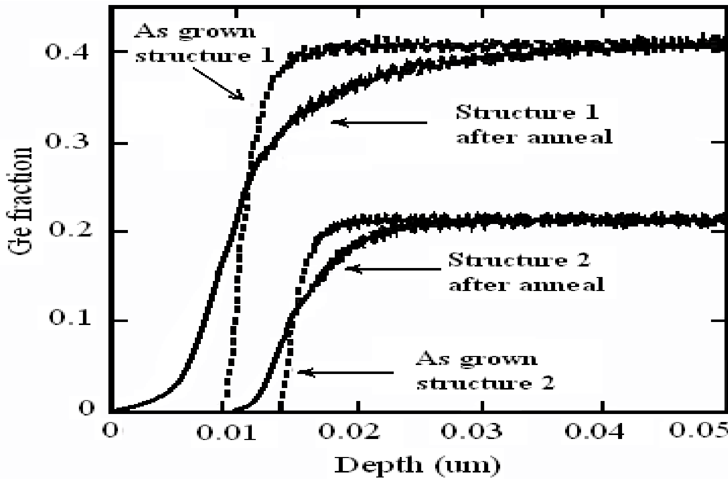



**FIGURE 5.4**
Ge profiles measured by SIMS for Boltzmann-Matano analysis. Both samples are annealed at 920$^o$C or 60 min in nitrogen. Source: MIT Annual Report 2005.

The Ge diffusion into the strained-Si layer from the $Si_{0.7}Ge_{0.3}$ layer can be formulated as a solution of Fick's second law for a pair of semi-infinite solids. The concentration c at a distance x from the initial interface at time t is given by

$$c = \frac{c_o}{2}\left[1 - erf\left(\frac{x}{2\sqrt{Dt}}\right)\right] \tag{5.1}$$

where D is a diffusion coefficient and $c_o$ is the initial Ge concentration in the $Si_{0.7}Ge_{0.3}$ layer. The Ge depth profiles obtained by SIMS were fitted with this equation and are shown as curves in Fig. 5.3(b). The measured points could be well fitted in the region where the Ge concentration was less than 5%. The discrepancies between the measured and calculated values in the high Ge-concentration region (> 5%) were due to a change in the Ge diffusion coefficient.

One issue for strained-Si/SiGe structures is the interdiffusion at the strained-Si/SiGe interface during device processing, which degrades device performance by reducing strain and carrier confinement and increasing alloy scattering. Research on Ge diffusion has been focused on Ge self-diffusion in $Si_{1-x}Ge_x$ and interdiffusion in compressively strained-SiGe structures with low Ge fractions (x < 0.25) grown on Si [13]. Basic understanding of interdiffusion, such as Ge fraction, strain, and temperature dependence, and the influence of point defects generated during oxidation, nitridation, and other processes is inadequate. In addition, little data is available for SiGe inter-

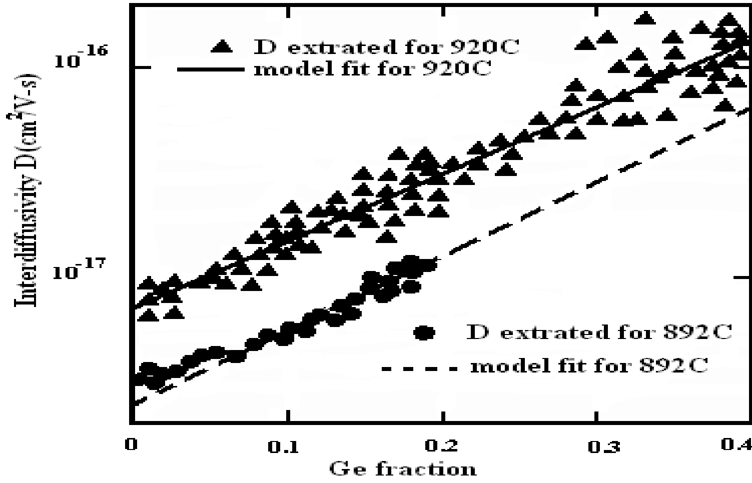



**FIGURE 5.5**
SiGe interdiffusivity (D) extracted from Ge profiles similar to those in Fig. 5.4 (symbols), and model fit (lines) at two different temperatures. Source: MIT Annual Report 2005.

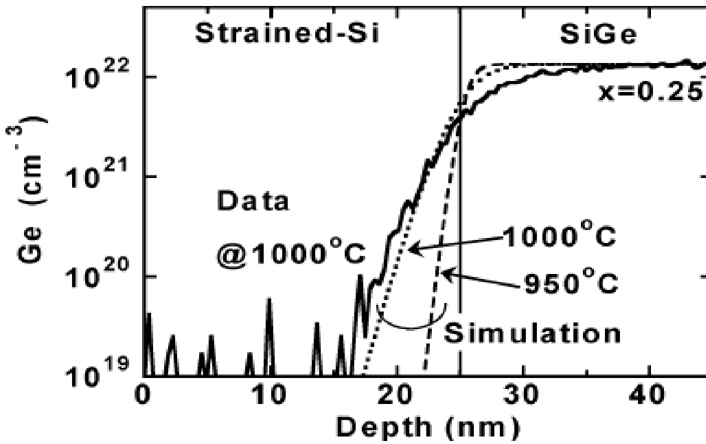



**FIGURE 5.6**
SIMS analysis for Ge atom profile (solid line) of strained-Si-SiGe layers in a strained-SOI. The Ge atoms diffuse into the strained-Si layer from the SiGe layers, although the simulation data (dashed line) at a 950$^o$C RTA process shows the suppression of the Ge diffusion. After T. Mizuno et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(1114-1121). With permission.

diffusion in device structures, such as strained-Si/strained-$Si_{1-y}Ge_y$/relaxed-$Si_{1-x}Ge_x$ with Ge content y > 0.3. Boltzmann-Matano analysis is used to obtain SiGe interdiffusivity experimentally. Fig. 5.4 shows an example of Ge profile used for Boltzmann-Matano analysis, for Ge fraction up to 0.4. The interdiffusivity model is based on the interdiffusion coefficients obtained from Boltzmann-Matano analysis (Fig. 5.5), which has a similar form to the model used in reference [13] for compressive SiGe at lower Ge mole fractions.

Fig. 5.6 shows the results of the secondary ion mass spectroscopy (SIMS) analysis (solid line) of the Ge profile of the fabricated strained-SOI structures. It was found that the Ge atoms diffuse to the strained-Si layer from the SiGe layer during the high temperature ($1000^oC$) annealing process. As a result, the strained-Si layer becomes thinner by 4 nm when the SiGe layer is defined as the region including Ge content of higher than 1%. On the other hand, the simulation results (dashed line) of the Ge distribution at the low temperature ($950^oC$) RTA process, the Ge diffusion length can be suppressed to only 1 nm. Therefore, the low temperature process is definitely required to suppress the Ge diffusion. The influence of the Ge diffusion on the device performance is discussed in chapter 7.

### 5.2.2 Kinetics: Oxidation of $Si_{1-x}Ge_x$ Layers

The microstructure of oxide layers on $Si_{1-x}Ge_x$ after atmospheric oxidation and the kinetics of the process have been described in several studies [14, 15, 16, 17, 18, 19, 20, 21]. During atmospheric pressure oxidation of Si, it is generally accepted that Ge acts as a catalyst for the formation of $SiO_2$ (i.e., it is neither consumed nor incorporated into the oxide). This catalytic effect causes the linear reaction constant in the well-known Deal-Grove oxidation model to be increased by a factor of 2-3. In the diffusion-limited parabolic regime, the presence of the Ge is reported to have no effect [16]. In cases of atmospheric pressure oxidation, Ge has been reported to accumulate at the oxide/semiconductor interface. The accumulation of Ge at the oxide/alloy interface was shown by capacitance-voltage measurements to produce a high density of fixed charge and a high interface trap density even after a forming gas anneal. However, the origin of this accumulation is not yet clear. Based on these early results it appears that atmospheric oxidation of $Si_{1-x}Ge_x$ alloys is limited for electronic passivation in applications such as metal oxide semiconductor devices.

High-pressure oxidation holds great promise for device quality thermal passivation of $Si_{1-x}Ge_x$ since it allows low temperature thermal passivation. This approach has been proven useful for both the oxidation of Si and Ge and, hence, its application to the alloys of $Si_{1-x}Ge_x$ is a natural step that has been only superficially explored. One important benefit of high-pressure oxidation has been shown to be the incorporation of Ge into the oxide during oxidation, which reduces or eliminates Ge accumulation at the interface.

Moderately high-pressure ($\sim 200$ atm) oxidation of alloys of dilute $Si_{1-x}Ge_x$

and superlattice structures at 740$^o$C has shown to produce oxides in which Ge was reportedly trapped in the oxide, though its state (chemical incorporation vs. entrapped elemental particles) has not yet been established. At much higher pressures (680 atm) and lower temperatures (550$^o$C) Ge was reported to be chemically incorporated into the oxide. Paine et al. [19] have compared high-pressure, low-temperature oxidation with conventional wet oxidation and explained the results in terms of the kinetics and thermodynamics of the oxidation process.

Strained-Si heterostructure substrates typically include a relatively thin (e.g., less than 250 Å) strained silicon layer that may be used as the channel in a MOSFET device. If the layer of strained-Si is grown too thick, misfit dislocation defects will occur in the layer, compromising the yield when MOSFET circuits are fabricated on the substrate. In particular, at a critical thickness, dislocations are favored for strain relief of the epitaxial film over continued accumulation of strain energy. The critical thickness is a function of the lattice mismatch between the epitaxial film and substrate, as well as the materials properties of both the epitaxial layer and the substrate. It is this critical thickness that may limit the useful strained-Si film thickness to less than, e.g., 250 Å.

Conventional oxidation techniques for thermally growing oxide layers on a Si substrate involve the consumption of a significant thickness ($\sim 0.44 t_{ox}$) of the Si substrate. If too much of the strained-Si layer is consumed in the oxidation process, then the layer will be too thin to obtain the benefits of the enhanced electron mobility. The minimum strained-Si film thickness required for significant mobility enhancement is approximately 50 Å [22]. Conventional methods of forming multiple gate oxides do not work well on a strained Si substrate since the strained-Si cap layer may be too thin to support the formation of both thick and thin gate oxides. This is particularly the case since during a typical MOSFET fabrication process, there is much additional strained-Si consumption due to various other processing steps (cleaning, thermal oxidation, and annealing). Therefore, forming gate oxide on strained-Si heterostructure substrates with the required gate oxide thicknesses but without sacrificing the enhanced electron mobility is a challenging task.

### 5.2.3 Oxidation of Strained-Si Layers

One of the key considerations in fabricating strained-Si channel devices is the epitaxial film quality, since its robustness is important for the performance enhancement and relies on the integrity of the strained layer structure. This is especially important in SiGe-MOS technology, where a low thermal budget process is necessary compared with other technologies, such as bipolar and BiCMOS. Strained-Si on relaxed-SiGe on insulator (SGOI) is very promising as it combines the benefits of two advanced technologies: (a) the conventional silicon on insulator and (b) the strained-Si. The SOI configuration offers various advantages associated with the insulating substrate, namely, reduced

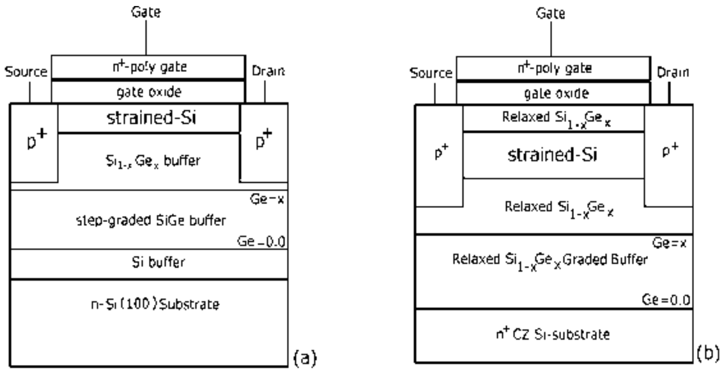



**FIGURE 5.7**
Schematic diagrams of (a) surface and (b) buried channel strained-Si MOSFETs.

parasitic capacitances, improved isolation, and reduced short-channel effect. High mobility strained-Si, strained-SiGe and strained-Ge channel MOSFETs have been reported on SOI substrates [1].

Silicon dioxide plays an important role in Si-integrated circuits where it is used for various purposes. Conventional oxidation techniques for thermally grown oxide layers on a Si substrate typically involve the consumption of a significant thickness of the Si substrate during the fabrication process. Also, there is an additional Si consumption due to various other processing steps such as cleaning and annealing. Strained-Si substrates include a relatively thin ($\sim$ 200 Å) strained-Si layer that may be used as the channel in a MOSFET (see Fig. 5.7). If too much of the strained-Si layer is consumed in the oxidation process, then the channel layer will be very thin to obtain the benefits of the enhanced electron mobility. The minimum strained-Si film thickness required for a significant mobility enhancement is $\sim$ 50 Å [23]. Also, conventional methods of forming multiple gate oxides do not work well on a strained-Si substrate since the strained-Si layer is very thin to support the formation of both thick and thin gate oxides. Therefore, the need to study the strained-Si oxidation has many motivations, including to have a high-quality gate oxide layer for MOSFET applications.

The thermal oxide on a Si substrate shows, in general, very good electrical properties. Unlike the Si, however, the thermal oxidation of strained-Si/SiGe films results in poor oxide quality and the Ge segregation during oxidation degrades the film quality [24, 25]. When strained-Si/SiGe/Si heterolayers are processed at a high temperature, beyond the film growth temperature, or during ion implantation the strain in the epilayers may relax and misfit dislocations are formed which degrade the device performance. In addition, after the gate oxidation, during gate polysilicon deposition and the dopant

activation, the diffusion of Ge into the $SiO_2$ gate takes place, which also causes device degradation.

Ultrathin $SiO_2$ gate dielectrics, of less than 20 Å in thickness, are needed for the 65-nm technology node [26]. Such a thin $SiO_2$ layer exhibits a significant leakage current ($> 1$ A/cm$^2$ at 2 V). In order to reduce the leakage current, an extensive search for alternative high dielectric constant gate dielectrics that would probably replace the $SiO_2$ for the sub-65 nm CMOS technologies is being pursued. Unlike the high temperature needed for the growth of $SiO_2$, most high-k dielectric deposition on Si can be made at a low temperature. Despite this fact, there are not many reports available in the literature on these dielectrics on strained-Si or strained-SiGe layers.

## 5.3 Rapid Thermal Oxidation

As discussed before, there is a strong evidence of oxidation-induced partial strain relaxation in relatively thick strained-Si/SiGe films which continue to increase as a function of oxidation temperature. It has also been reported in the literature that the mobility of strained-Si/SiGe degrades due to the presence of high fixed oxide charge ($\sim 4\times10^{11}$ cm$^{-2}$) [6, 7, 8]. For the fabrication of buried-channel MOSFETs (see Fig. 5.7) in SiGe-technology, it is essential to form ultrathin gate dielectrics directly on strained-SiGe layers. Unfortunately, for gate dielectric formation, the conventional high temperature thermal oxidation (and high thermal budget) process cannot be applied to strained-Si/SiGe layers due to strain relaxation, dopant and Ge segregation/diffusion, and propagation of dislocations as discussed before. Due to its low thermal budget and high thickness controllability for ultrathin dielectrics, rapid thermal oxidation has become very popular for dielectric formation on strained-Si/SiGe layers [27].

To enhance the reliability of ultrathin dielectrics required for MOS devices, films with low leakage current, improved charge trapping characteristics, and high dielectric strength are necessary. To suppress charge trapping and improve the reliability of gate dielectrics, incorporation of nitrogen at the Si/$SiO_2$ interface have been proposed [28, 29]. Reoxidized nitrided oxide (RNO) and thermally nitrided (in $NH_3$) oxides have also shown excellent characteristics toward hot carrier effects with reduced interface state generation and act as a barrier toward impurity penetration into $SiO_2$. However, electron trapping is inherent in nitrided oxide due to hydrogen, and the reoxidation step is needed in reducing the H atoms [30, 31]. Rapid thermal processing in a nitrogen ambient ($N_2O$ or NO) is essentially free of H atoms and results in oxides that are highly resistant to hot carrier effects with reduced electron trapping [32, 33, 34].

Oxidation of strained-Si films on relaxed-$Si_{0.75}Ge_{0.25}$ layer using a rapid thermal technique in $N_2O$ ambient has been reported by Tan et al. [35]. The samples were cleaned by piranha ($H_2SO_4$:$H_2O_2$ = 3:1) instead of a normal RCA/SC1 cleaning step ($H_2O$:$H_2O_2$:$NH_4OH$ = 5:1:1) which is known to etch the thin strained-Si epilayer. Prior to loading the samples into the oxidation chamber, they were dipped for 10 min in diluted HF (48%HF:$H_2O$ = 1:50) leaving the strained-Si surface both hydrophobic and hydrogenated. The ramp-up rate of the rapid thermal processor (AST SHS-10) is 30$^o$C/s while the ramp-down rate is lower, which leads to a reduced thermal exposure for the strained-Si and also a better control of thin oxide thickness. RTO was performed at 1000$^o$C for 400 s in $N_2O$ ambient. The oxide thickness was found to be ~ 90 Å from an ellipsometric thickness measurement. Immediately after oxidation, 250 nm of Al was evaporated on top of the oxide to form MOS capacitors with an area of $7.85\times10^{-3}$ $cm^{-2}$. Ohmic contact was formed at the back with Al. The MOS capacitors were annealed in forming gas ($H_2$:$N_2$ = 1:9) at 400$^o$C for 5 minutes after metallization.

It has been shown that strained-Si on SiGe layers are stable with respect to high-temperature processing steps with no mobility reduction for thermal treatments up to only 800$^o$C for 30 min and 950$^o$C for 3 min [36]. Also, simulation of the Ge diffusion in Si/SiGe heterostructures, indicates that the degradation of the heterostructures takes place after high-temperature processing steps, which may be attributed to the movement of Ge into the strained-Si channel. A simple model was used to estimate the amount of Ge outdiffusion into the strained-Si epilayer. The Ge concentration at a particular temperature, "c", as a function of time and distance, is given by [36]

$$c(x,t) = \frac{c_o}{2}\left[erf\left(\frac{x+\frac{d}{2}}{2\sqrt{Dt}}\right) - erf\left(\frac{x+\frac{d}{2}}{2\sqrt{Dt}}\right)\right] \qquad (5.2)$$

where D is the diffusivity of Ge in Si and is given by D = $D_o$ exp(-$E_a$/kT) with $D_o$ = 400 $cm^2s^{-1}$ and $E_a$ = 4.7 eV. The initial Ge concentration in the $Si_{0.75}Ge_{0.25}$ cap layer ($c_o$) is 25% and the $Si_{0.75}Ge_{0.25}$ cap layer thickness (d) is 5 $\mu$m. Note that x = 0 corresponds to the center of the $Si_{0.75}Ge_{0.25}$ cap layer.

A simulation was done to model the extent of Ge diffusion within the heterostructure under thermal stress, assuming that Ge can diffuse infinitely [35]. For the simulation, an initially perfect Si/$Si_{0.75}Ge_{0.25}$ heterointerface was assumed, and a one-dimensional representation of Fick's law was used to calculate the diffusion of Ge atoms from the Si/$Si_{0.75}Ge_{0.25}$ layer into the Si channel under different anneal conditions. However, in the simulation the effect of elastic strain, crystal dislocations, and various other material conditions were not included. Simulation results are shown in Fig. 5.14, which shows the Ge profile at the strained-Si/relaxed-SiGe interface before and after oxidation. It is seen that a significant amount of Ge atoms have outdiffused into the strained-Si layer under the present oxidation condition (1000$^o$C for

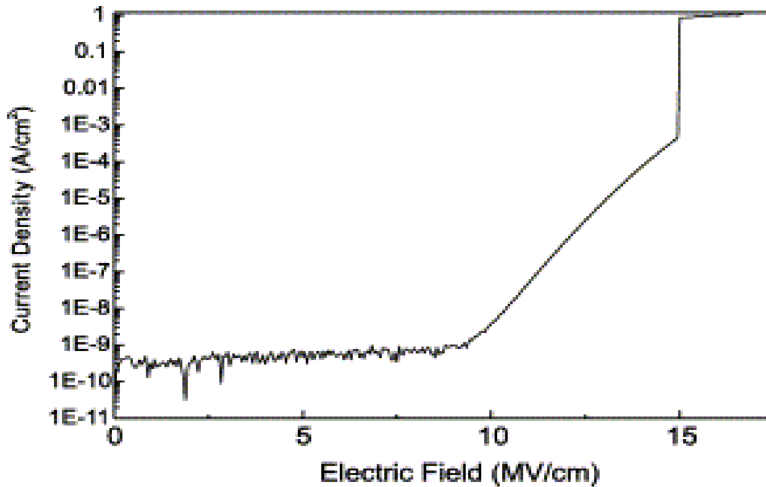



**FIGURE 5.8**
J-E characteristic of a MOS capacitor. F-N tunneling behavior is clearly seen for an electric field higher than 9 MV/cm and oxide breakdown at 15 MV/cm. After C. S. Tan et al., *Solid-State Electron.*, Vol. 45, 2001(1945-1949). With permission.

400s). This gives rise to a high interface states density and a negative fixed charge at the $SiO_2$/strained-Si interface.

To study the electrical properties, the J-E characteristic of the MOS capacitor, biased into accumulation, was measured using HP-4061A measurement system and is shown in Fig. 5.8. In the plot of I vs. E in Fig. 5.8, there are two distinct regions of $E$, namely, $< 9$ MV/cm and $> 9$ MV/cm, where two different leakage current characteristics can be identified. A relatively high leakage current, observed in the $> 9$ MV/cm region, shows a strong dependence on the oxide field and is an indication of F-N tunneling, as is observed in most of the dielectrics at high fields. To determine the dominant leakage current mechanism in the $SiO_xN_y$ oxide a plot of $\ln(J/E^2)$ vs. $1/E$ (F-N mechanism) has been used. $SiO_xN_y$ gate oxides show a linear relationship and the conduction mechanism is dominated by the P-F mechanism at a low field.

An analysis on the F-N current gives an electron effective mass of $0.45m_o$ ($m_o = 9.1\times10^{-31}$ kg) and a barrier height of $\sim 3.0$ eV. The oxide is also found to have a breakdown field as high as 15 MV/cm. This high breakdown field is due to the nature of $N_2O$ oxidation [34]. For electric field below 9 MV/cm, the leakage current is higher than that predicted by the F-N tunneling mechanism (see Fig. 5.8). A high leakage current may be attributed to the shallower traps lying closer to the conduction band edge, which generate a hopping current.

In order to enable operation at the ultrashort channel lengths, ultrathin

body and double-gate transistors are being investigated as alternate structures for future silicon MOSFETs [26]. The ability to combine the advantages of alternate device structures with strained-Si offers many potential advantages such as reduction of channel doping compared to bulk MOSFETs. Rapid thermal annealing (RTA) of the strained-Si on SOI structures has also been carried out at different temperatures [37]. Thin (130 Å) strained-Si films grown directly on an insulator (SiGe-free) and strained-Si on insulator with an intermediate SiGe layer were subjected to the following annealing conditions in nitrogen ambient: $600^o$C for 10 s, $800^o$C for 10 s, and $950^o$C for 1 s. For the strained-Si, directly on insulator, a single strained-Si peak was observed at 512 $cm^{-1}$ while the control-Si reference peak was found at 521 $cm^{-1}$. For the strained-Si with underlying SiGe buffer, two peaks were observed: a highly attenuated SiGe peak at 506 $cm^{-1}$, and a strained-Si peak at 512 $cm^{-1}$. Thus the Raman data confirms that the underlying SiGe buffer layer is not necessary to maintain the strain in thin strained-Si films on an insulator (SSOI). No strain relaxation was also observed following a rapid thermal annealing of SSOI layers to a temperature as high as $950^o$C.

## 5.4 Plasma Nitridation of Strained-Si

For strained-Si channel MOSFET fabrication, conventional thermal oxidation at temperatures ranging from 700 to $850^o$C has been employed [5, 6, 8]. In order to form the gate oxide for a hetero-MOSFET (see Figs. 5.7(a) and (b)), the strained-Si layer needs to be oxidized in case of surface-channel device, while the top SiGe layer need to be oxidized directly in case of the buried-channel device. Unfortunately, there are problems due to the nature of the Si/SiGe heterostructures in both cases that render the direct oxidation process unsatisfactory.

During the last two decades, many attempts have been made both to quantify the effect, in terms of Ge concentration at the interface, temperature and ambient (wet vs. dry), and to use the conventional Deal-Grove model to study the oxide growth kinetics. There is now a good deal of understanding of the kinetics of thermal oxidation of SiGe layers and also the multilayer structures of the grown oxides [38]. However, the main conclusion of these studies is that the temperature normally used for thermal Si oxidation is too high to preserve the crystalline perfection of the Si/strained-SiGe layers. In order to circumvent this problem, different low-temperature oxidation methods such as ECR/microwave plasma oxidation, ion implantation, and photon (mainly UV) assisted oxidation have been proposed [1]. In most of these techniques, a process temperature, lower than $400^o$C, is used for a typical 10 to 50-nm thick oxides without any significant relaxation of the strained

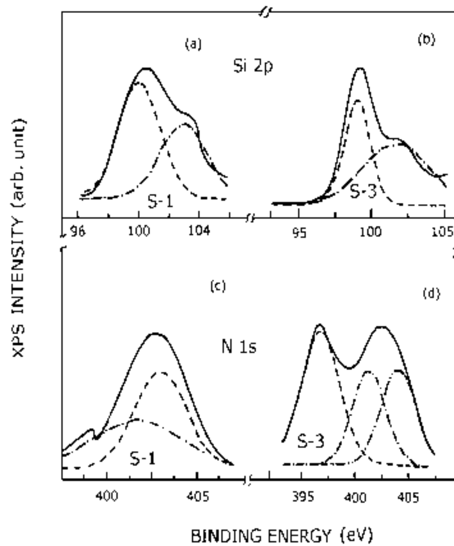



**FIGURE 5.9**
Deconvolution of Si 2p spectra obtained by high-resolution XPS of 1 min $Ar^+$ ion etched: (a) $O_2/N_2O$ nitrided oxide (S-1) and (b) $O_2/NH_3/N_2O$ nitrided oxide (S-3), (c) N 1s spectra of 1 min $Ar^+$ ion etched $O_2/N_2O$ nitrided oxide (S-1), and (d) $O_2/NH_3/N_2O$ nitrided oxide (S-3). After L. K. Bera, *Studies on Applications of Strained-Si for Heterostructure Field Effect Transistors (HFETs)*, Ph.D. Thesis, IIT Kharagpur, 1998.

layers. To achieve a high-quality gate oxide and still maintain the strain in the film, low-temperature plasma oxidation of strained-Si/SiGe films has also been reported by several authors [39, 40].

Long-term reliability, electrical stability, and poor diffusion barrier properties of sub-100 Å thick oxides have created interest in alternative dielectrics to silicon dioxide. Nitroxides and reoxidized nitroxides are very promising for gate dielectric applications. Nitroxides have shown good diffusion barriers against impurities and superior hot-electron immunity compared to thermal oxides [41]. In the following, we examine the nitridation conditions to find a good starting nitroxide for device applications and the effect of annealing of strained-Si films, grown at 700$^o$C on a fully relaxed step-graded SiGe buffer layer using a GSMBE, in $N_2O$- and $NH_3$-plasma [11].

In a typical growth process, the microwave plasma system was first evacuated to $10^{-3}$ Torr, followed by the introduction of $O_2$, $N_2O$, and $NH_3$ gases into the discharge chamber. No external heating of the substrate was made;

**TABLE 5.1**
Processing conditions for nitridation of $SiO_2$ films on strained-Si.

| Samples | Oxidation | Nitridation | ROXNOX |
|---|---|---|---|
| S-1 | $O_2$ plasma 1 min. | $N_2O$ plasma 1 min. | - |
| S-2 | $O_2$ plasma 1 min. | $NH_3$ plasma 1 min. | - |
| S-3 | $O_2$ plasma 1 min. | $NH_3$ plasma 1 min. | $N_2O$ plasma 1 min. |

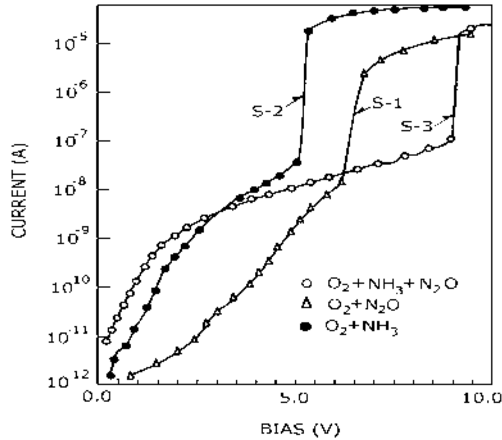



**FIGURE 5.10**
I-V characteristics of $O_2/N_2O$ nitrided oxide, $O_2/NH_3$ nitrided oxide, and $O_2/NH_3/N_2O$ nitrided oxide samples. After L. K. Bera, *Studies on Applications of Strained-Si for Heterostructure Field Effect Transistors (HFETs)*, Ph.D. Thesis, IIT Kharagpur, 1998.

however, the discharge itself raised the substrate temperature to about $150^oC$. Ultrathin $SiO_2$ films ($\sim$ 25Å, to be used as a seed layer) were grown in $O_2$ plasma. Three different samples (hereafter referred to as S-1, S-2, and S-3) were nitrided and the processing conditions are given in Table 5.1. Thicknesses of the nitrided $SiO_2$ films were in the range of 75-100 Å.

XPS (VG Scientific ESCALAB MK-II spectrometer equipped with a concentric hemispherical analyzer) with $AlK_\alpha$ radiation was used to profile the incorporation of nitrogen in the nitrided films and also to investigate the chemical structure of the films from the shift of the core-level binding energy. This technique is particularly attractive for deducing local atomic structure due to the sensitivity of the spectra to the chemical environment. The surfaces of (as-grown and $Ar^+$ ion etched samples) $SiO_2$, nitrided by $N_2O/NH_3$ plasma, were also analyzed using high-resolution XPS to investigate the modification of chemical bonding.

Fig. 5.9 shows the Si 2p, and N 1s photoemission spectra for 1 min $Ar^+$-ion

etched samples S-1 and S-3. No nitrogen signal was observed in the XPS spectra from the as-grown film surfaces. A chemical shift of $\sim$ 3.7 eV is observed for Si 2p peak with respect to that of bulk-Si. The observed binding energy is inbetween silicon nitride and silicon dioxide, indicating the formation of silicon oxynitride films. Since nitrogen is less electronegative compared to oxygen (electronegativity of 3.04, as compared to 3.44 for oxygen), Si–N bonds reduce the binding energy of core-level electron, resulting in a shift of Si 2p peak to a lower value.

The observed Si 2p peaks are attributed to signals from the Si substrate ($\sim$ 98.8 eV) and $SiO_xN_y$ layer (around $\sim$ 101-103 eV). Deconvolution of Si 2p spectra (Figs. 5.9(a) and (b)) confirmed that peaks at lower energy ($\sim$ 98.8 eV) originate from the strained-Si (substrate) and that of higher energy ($\sim$ 101-103) from interfacial $SiO_xN_y$ layer. N 1s spectra for $N_2O$ nitrided $SiO_2$ (S-1) show a single peak at 403.0 eV, originating from N–O bonds. The major contribution of N–O bonds is clearly observed from the deconvoluted N 1s spectra as shown in Fig. 5.9(c) which consists of Si $\equiv$ N and $Si_2$=N–O bonds. It is also noticed that the peak intensity of $Si_2$=N–O bond is higher than Si $\equiv$ N bond. On the other hand, N 1s spectra of the $N_2O$-plasma treated $NH_3$-nitrided $SiO_2$ films (S-3) exhibit two peaks. The first peak corresponds to Si $\equiv$ N bond at 396.8 eV and the convoluted spectra of the second peak (at 403.7 eV) with Si–N=$H_2$ and $Si_2$=N–O bonds are shown in Fig. 5.9(d).

It may be noted that the $N_2$ atoms accumulate at or near the oxide/strained-Si interface forming a silicon oxynitride layer, which is believed to provide a greater immunity to electrical stress. Similar observation has also been reported in case of bulk-Si [42, 43]. The I-V characteristics of MIS capacitors on all the samples, measured under a negative gate bias condition, are shown in Fig. 5.10. Although sample S-3 exhibits a higher low-field leakage current (compared to S-1 and S-2), it shows a higher breakdown voltage of about 9 V. This may be due to the enhancement of nitrogen incorporation at $SiO_2$/Si interface to form the oxynitride, which has a narrow bandgap. As a result, the barrier height is lowered which increases the current density at low electric field. A typical resistivity value of sample S-3 is found to be about $2.7\times10^{15}$ $\Omega$-cm.

High-frequency (1 MHz) capacitance-voltage (C-V), current-voltage (I-V) and conductance-voltage (G-V) characteristics of the Al-gate MIS capacitors (area $1.96\times10^{-3}$ $cm^2$) fabricated using the nitrided films were used to study the electrical properties and trapping behavior. F-N injections on capacitors were performed at a constant current level of 10 mA/$cm^2$ (gate injection) which correspond to an average electric field in the oxide of 9-10 MV/cm. C-V and G-V characteristics were used to determine the flatband voltage and interface state density.

Using C-V and G-V characteristics, fixed insulator charge density for samples S-1, S-2, and S-3 are found to be $1.5\times10^{11}$ $cm^{-2}$, $2.0\times10^{11}$ $cm^{-2}$, and $1.2\times10^{11}$ $cm^{-2}$, respectively. A low value of $Q_f/q$ for S-3 may be due to outdiffusion of $H_2$ during reoxidation and reduction of strained bond by enhanced

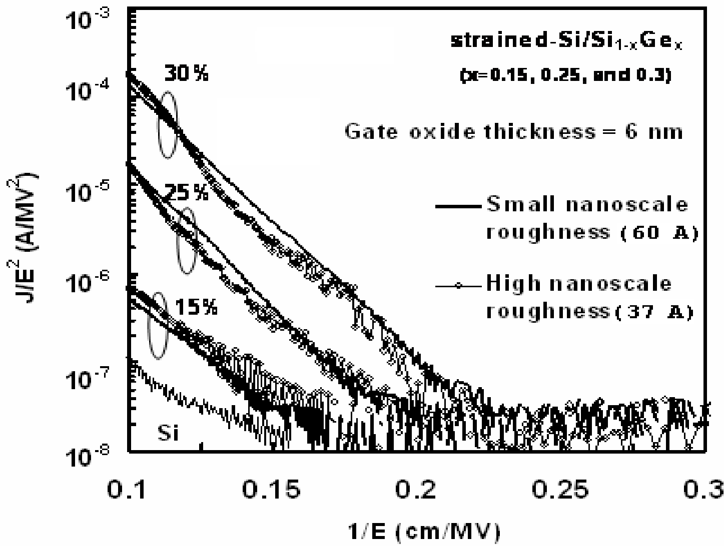



**FIGURE 5.11**
Effect of cross-hatch roughness on F-N tunneling characteristics. After K. S. K. Kwa et al., EMC Conference 2005.

nitrogen incorporation at strained-Si/$SiO_2$ interface. The interface trap density at midgap, computed using Hill's method [44] for samples S-1, S-2, and S-3 were found to be $3.3\times10^{11}$ cm$^{-2}$eV$^{-1}$, $5.8\times10^{11}$ cm$^{-2}$eV$^{-1}$ and $3.5\times10^{11}$ cm$^{-2}$eV$^{-1}$, respectively.

## 5.5 Effect of Surface Roughness

As discussed earlier, quantitative topographic measurements show that the surface roughness of the strained-Si layer is associated with the cross-hatch patterns, which basically result from a local accumulation of the underlying misfit dislocations beneath the surface. Evidence of the variation in the oxidation rate of strained-Si along the cross-hatch has been reported by Olsen et al. [45, 46]. The undulating oxide thickness was found to be accompanied by increased nanoscale roughness at the Si/$SiO_2$ interface. The impact of cross-hatching on the performance of strained-Si surface channel devices has also been investigated [47]. As shown in Fig. 5.11, the critical electric field at which FN tunneling begins increases significantly (from 4 MV/cm to 7 MV/cm) as cross-hatching roughness is reduced from 8 nm to 4 nm. The authors fabri-

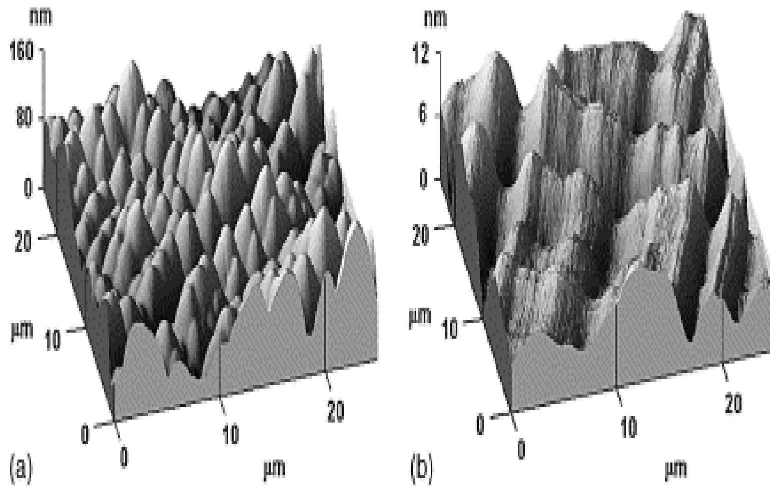



**FIGURE 5.12**
AFM images of as-grown Si/SiGe material grown by: (a) GSMBE and (b) ULPCVD. The dissimilar scales highlight the difference in morphology due to growth technique. After S. H. Olsen et al., *Solid-State Electron.*, Vol. 47, 2003(1289-1295). With permission.

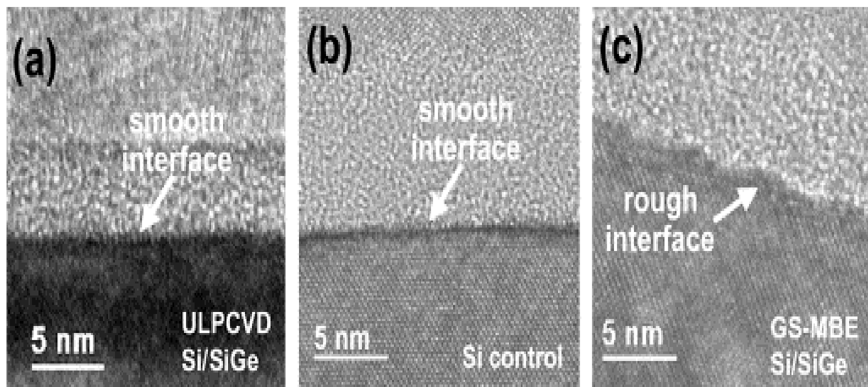



**FIGURE 5.13**
HRTEM images of the $Si/SiO_2$ interface roughness underneath gate structures on fully processed wafers: (a) ULPCVD-grown films at a $0^o$ vicinal angle, (b) control-Si at a $0^o$ vicinal angle, (c) GSMBE films at a $3^o$ vicinal angle. After S. H. Olsen et al., *Solid-State Electron.*, Vol. 47, 2003(1289-1295). With permission.

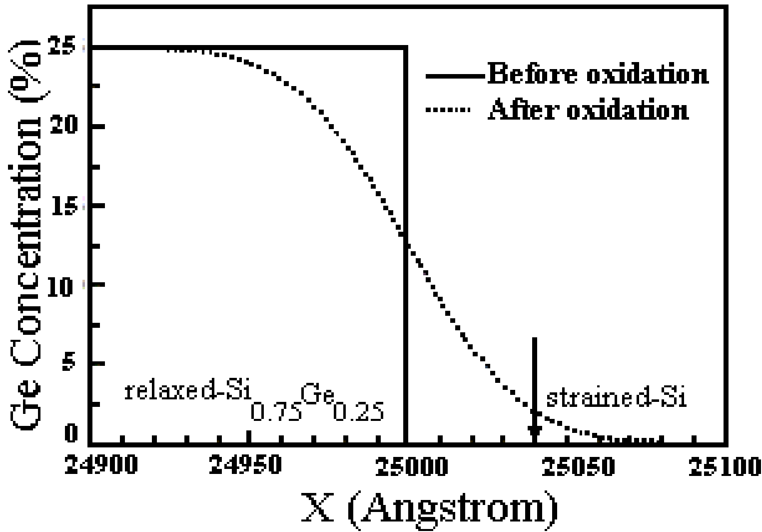



**FIGURE 5.14**
Ge profile at the strained-Si/SiGe interface before and after rapid thermal oxidation at 1000°C for 400 s in $N_2O$ ambient. After C. S. Tan et al., *Solid-State Electron.*, Vol. 45, 2001(1945-1949). With permission.

cated strained-Si surface channel devices using a standard MOS process with a reduced thermal budget on Si/SiGe layers grown by both GSMBE and ultra-low-pressure chemical vapor-deposited strained-Si films. The average cross-hatch roughness was found to be more than an order of magnitude lower on ULPCVD-grown films compared with GSMBE-grown films, as shown in Fig. 5.12.

The critical strained-Si/$SiO_2$ interface roughness has been shown to be strongly dependent on the vicinal angle. Smooth oxide interfaces were evident in regions where the angle is parallel to (001)Si, whereas distinctive roughness was displayed below oxide grown on the sloping edges of the cross-hatch pattern. High-resolution TEM (HRTEM) images of the strained-Si/$SiO_2$ interfaces at high and low vicinal angles are shown in Fig. 5.13. HRTEM analysis carried out at regions of the GSMBE strained-Si/$SiO_2$ interface on higher vicinal angles showed that increased roughness accompanied the sloping edges of the cross-hatch pattern.

The relationship between vicinal angle (and thus cross-hatch roughness) and gate oxide interface roughness identified on several strained-Si films is summarized in Table 5.2. The authors concluded that the average mobility degradation due to oxide interface scattering on devices fabricated on ULPCVD-grown films is expected to be greatly suppressed compared with devices fabricated on GSMBE films. Peak-to-peak roughness measurements

**TABLE 5.2**
Summary of gate oxide interface roughness measured on device structures fabricated at various vicinal angles on strained-Si/SiGe films using ULPCVD, GSMBE and unstrained bulk-Si(100). After S. H. Olsen et al., *Solid-State Electron.*, Vol. 47, 2003(1289-1295). With permission.

| Film/Vicinal Angle at roughness | $Si/SiO_2$ interface roughness (nm) | measurement location ($^o$) |
|---|---|---|
| bulk-Si | 0 | 0.20 |
| ULPCVD strained-Si/SiGe | 0 | 0.28 |
| GSMBE strained-Si/SiGe | 0 | 0.27 |
| GSMBE strained-Si/SiGe | 3 | 0.55 |

carried out using the AFM indicate that vicinal angles as large as $7^o$ are present on GSMBE material and additional increases in the $Si/SiO_2$ interface roughness in these areas are anticipated.

## 5.6 Effect of Strained-Si Layer Thickness

Strained-Si devices have also been shown to be scalable [48] and n-MOSFETs with gate lengths down to 25 nm have shown 15-30% performance enhancement over bulk-Si. Strained-Si technology is moving toward production. Several research groups have demonstrated both n- or p-MOSFETs with excellent characteristics fabricated with a wide variety of strained-Si layer thicknesses, many of which are greater than the strained-Si critical thickness. However, the effect on a MOSFET of the misfit dislocations that are generated by exceeding the critical thickness has not been thoroughly studied. It has been suspected that misfit dislocations may have a harmful effect on MOSFET characteristics, but their exact impact has not been precisely identified nor explained [9, 49]. Toward strained-Si technology development, new issues must be studied that are not yet fully examined. One of these issues is the limitation on the maximum strained-Si film thickness for a manufacturable MOSFET technology. Guidelines for the minimum strained silicon thickness have been reported [22, 49], but the constraints on the maximum film thickness have yet to be defined.

The presence of Ge in the strained-Si channel region can degrade electrical performance of the device has also been reported [50]. Effect of strained-Si layer thickness on the electrical properties of thermally grown gate oxide has been investigated by Bera et al. [51]. The authors have investigated the electrical properties of MOS capacitors grown on different thickness (80-500 Å) of strained-Si layers on 30% relaxed-SiGe buffer. Thermal oxidation

at 800$^o$C was carried out to form $\sim$ 50 Å thick gate oxide after channel implants and thermal activation. A 2000 Å thick polysilicon doped with B (by implantation) was used as gate electrode. Ti-silicide was formed on gate polysilicon as a part of self-aligned p-MOSFET fabrication. The Al/1%-Si was deposited in the backside of the wafers for substrate contact. The samples were annealed at 400$^o$C for 30 min in forming gas.

Fig. 5.15 shows the normalized quasistatic capacitance voltage characteristics of MOS capacitors on different strained-Si and bulk-Si wafers. The oxide thickness ($\sim$ 57Å) was estimated from the accumulation capacitance, which is in close agreement with value measured by ellipsometer. A plateau is observed in accumulation region for thin (80Å) strained-Si. This is due to the charge confinement in the strained-Si/SiGe heterostructure [10, 52]. The flat band voltage for 80 Å strained-Si is shifted toward positive bias, indicating the presence of negative fixed charge in the oxide. These negative fixed oxide charges originate from the presence of unsaturated Si-O- bonds in the oxide. In the case of thin strained-Si layer, high-temperature processing prior to gate oxidation might cause Ge diffusion into the film. During gate oxidation, diffused Ge species are oxidized and form Si-O-Ge bonds. It is known that Ge-O bonds are weak and can easily be broken into elemental Ge and Si-O- bonds [53]. Fig. 5.16 shows the calculated interface state density vs. energy characteristics for different samples. $D_{it}$ is higher for thin strained-Si compared to thicker strained-Si and bulk-Si samples. The higher interface state density for thin strained-Si is also due to the diffusion of Ge from the buffer layer.

Fig. 5.17 shows conductance-voltage characteristics of MOS capacitor on thin (80 Å) and thick (500 Å) strained-Si with frequency as a parameter. The conductance curve from thin (80 Å) strained-Si shows dispersion in two bias regions, A and B, corresponding to depletion and inversion bias, respectively. Capture and emission of majority carriers by interface traps causes an energy loss which is measured as an equivalent parallel conductance [54]. Therefore, two peaks in G-V characteristics indicate that two different types of traps are present in the oxide.

Fig. 5.18 shows the plot of natural log of carrier recombination time constant vs. normalized (by kT/q) surface potential. Both A and B regions show straight lines having opposite slopes, indicating two different types of states are distributed in the oxide/substrate interface. Since the same defect can act as one type of trap for electrons but as another type for holes, classification of traps is important. The concept of capture cross-section is the best way to classify the traps. The capture cross-section values for regions A and B, as determined by the ordinate at the origin are $1.3\times10^{-15}$ and $4.1\times10^{-21}$ cm$^2$, respectively. Thus, it is clear that conductance peaks in region A are due to neutral traps while those in region B are due to Coulombic repulsive-type traps. It is likely that the distribution of diffused Ge species is close to the strained-Si/SiGe interface and Ge do not reach $Si/SiO_2$ interface in the case of thick (500 Å) strained-Si layer. Ge species diffusing into the $Si/SiO_2$ interface causes very high coulombic repulsive types of trap density for thin

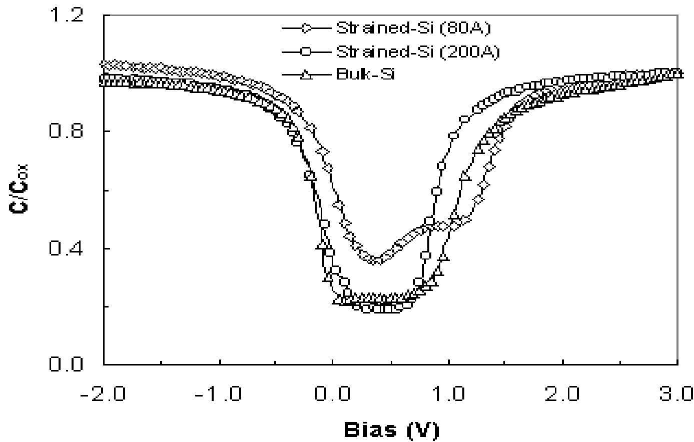



**FIGURE 5.15**
Quasistatic C-V characteristics of MOS capacitor on strained-Si and bulk-Si. After L. K. Bera et al., *Thin Solid Films*, Vol. 462-463, 2004(85-89). With permission.

samples. As mentioned above, Si-O- bonds can trap electrons and act as negative charge centers and the Si-O- bonds at the interface are responsible for the coulombic repulsive traps.

Fig. 5.19 shows the gate leakage current vs. gate bias characteristics of MOS capacitors on different thickness of strained-Si. F-N tunneling starts at higher voltages for thinner strained-Si sample, indicating the presence of a thicker gate oxide. The probable reasons are:

• Actual thickness of strained-Si layer reduces during different cleaning processes and also by gate oxide formation.

• The entire strained-Si layer may be consumed and the SiGe layer may be exposed to oxidation process, if the strained film is too thin.

• Ge diffusion occurs during thermal processes like implant activation, nitride deposition, and gate oxidation. For thicker strained-Si film, Ge diffusion from the buffer layer does not reach the surface. Presence of Ge in thin strained-Si film enhances oxidation rate [55].

The effect of constant voltage stressing of the oxide (Fig. 5.20) shows the stress-induced leakage current produced during constant voltage stressing of 6.0 V for 200 s. The charge trapping behavior, monitored during constant current stress under gate injection, as shown Fig. 5.21, indicates the trapping of negative charge in the oxide. At low current stressing, sinusoidal oscillations were found in the gate voltage signal (see Fig. 5.22). These fluctuations diminish at higher stress currents. The trapping-detrapping mechanism may be the cause of such gate voltage fluctuation at low stress levels.

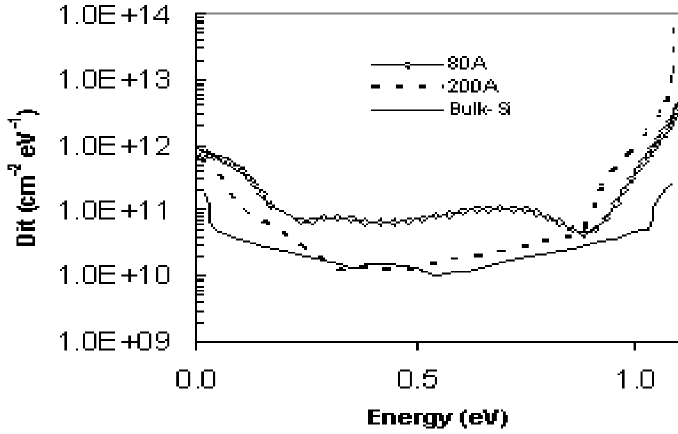



**FIGURE 5.16**
Interface state density ($D_{it}$) vs. energy of capacitors on strained-Si. $D_{it}$ increases with the decrease of strained-Si thickness. After L. K. Bera et al., *Thin Solid Films*, Vol. 462-463, 2004(85-89). With permission.

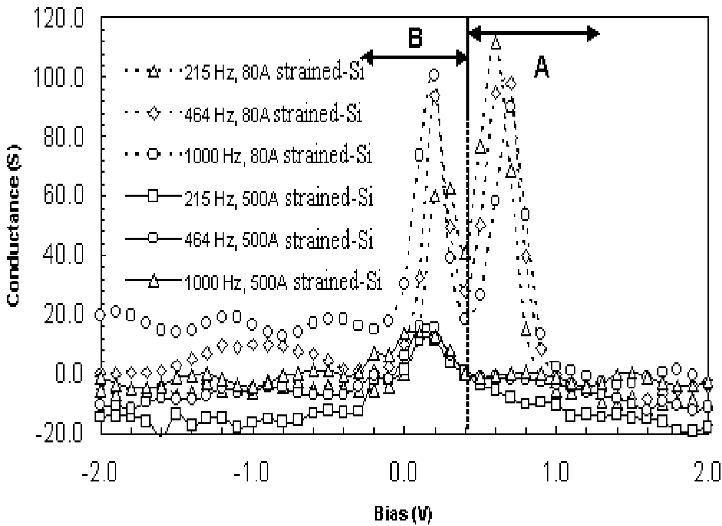



**FIGURE 5.17**
Conductance vs. gate bias characteristics of MOS capacitors on strained-Si. After L. K. Bera et al., *Thin Solid Films*, Vol. 462-463, 2004(85-89). With permission.

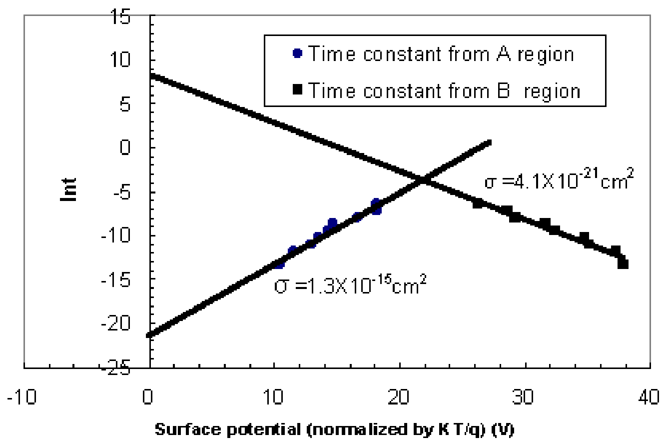



**FIGURE 5.18**
Experimental majority carrier interface recombination time constant vs. surface potential calculated from the region A and B (Fig. 5.17) measured at different frequencies. After L. K. Bera et al., *Thin Solid Films*, Vol. 462-463, 2004(85-89). With permission.

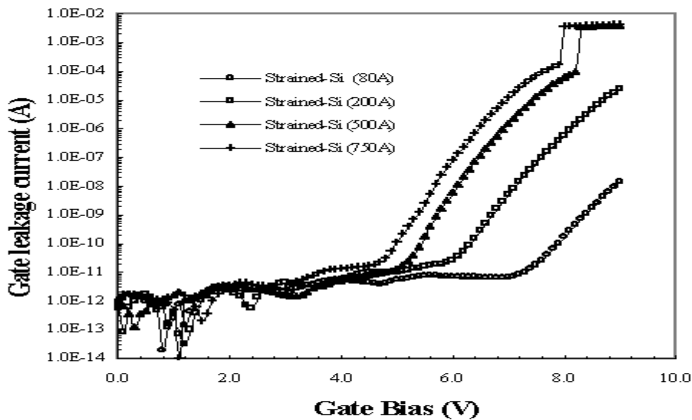



**FIGURE 5.19**
Gate current vs. gate bias (I-V) characteristics of capacitors on strained-Si. After L. K. Bera et al., *Thin Solid Films*, Vol. 462-463, 2004(85-89). With permission.

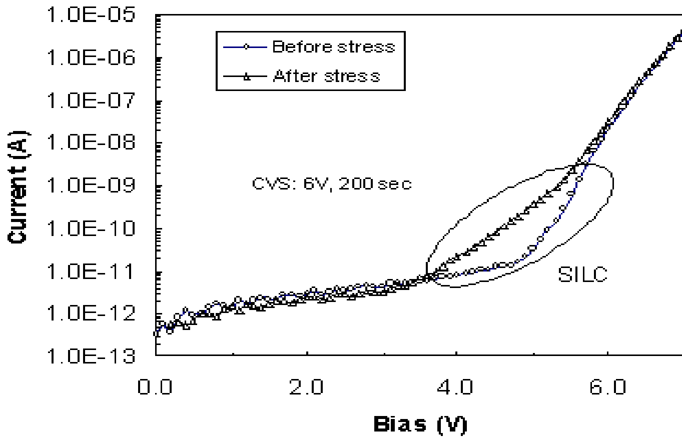



**FIGURE 5.20**
I-V characteristics of capacitors before and after constant voltage stress (+ 6 V) for 200 sec. After L. K. Bera et al., *Thin Solid Films*, Vol. 462-463, 2004(85-89). With permission.

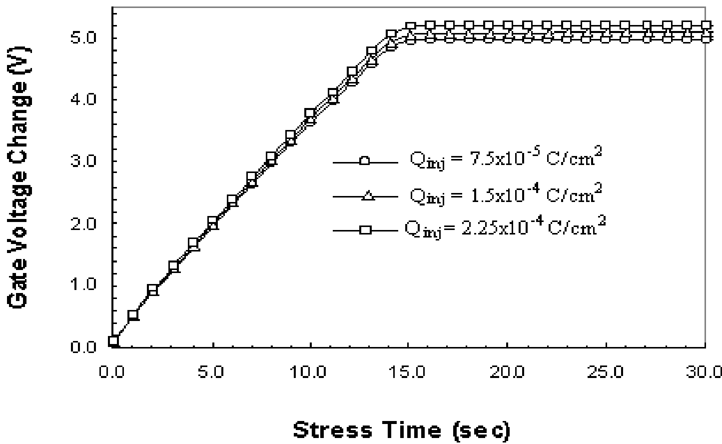



**FIGURE 5.21**
Gate voltage change vs. stress time during constant current stress on a MOS capacitor. After L. K. Bera et al., *Thin Solid Films*, Vol. 462-463, 2004(85-89). With permission.

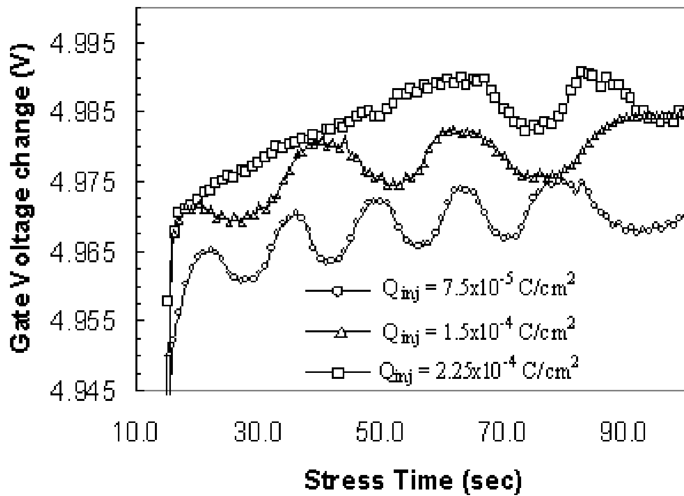



**FIGURE 5.22**
Oscillation in gate voltage shift during constant current stressing. After L. K. Bera et al., *Thin Solid Films*, Vol. 462-463, 2004(85-89). With permission.

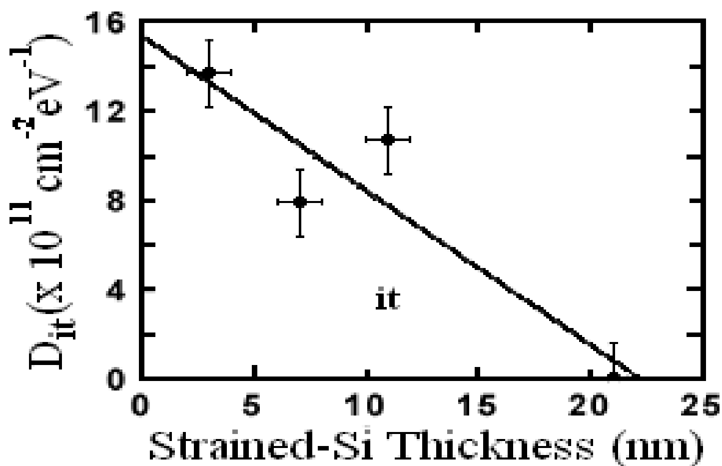



**FIGURE 5.23**
Interface state density of the gate oxide/strained-Si layer at room temperature determined from the subthreshold swing of p-MOSFETs. $D_{it}$ is found to increase rapidly with decreasing strained-Si thickness. After T. Mizuno et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(1114-1121). With permission.

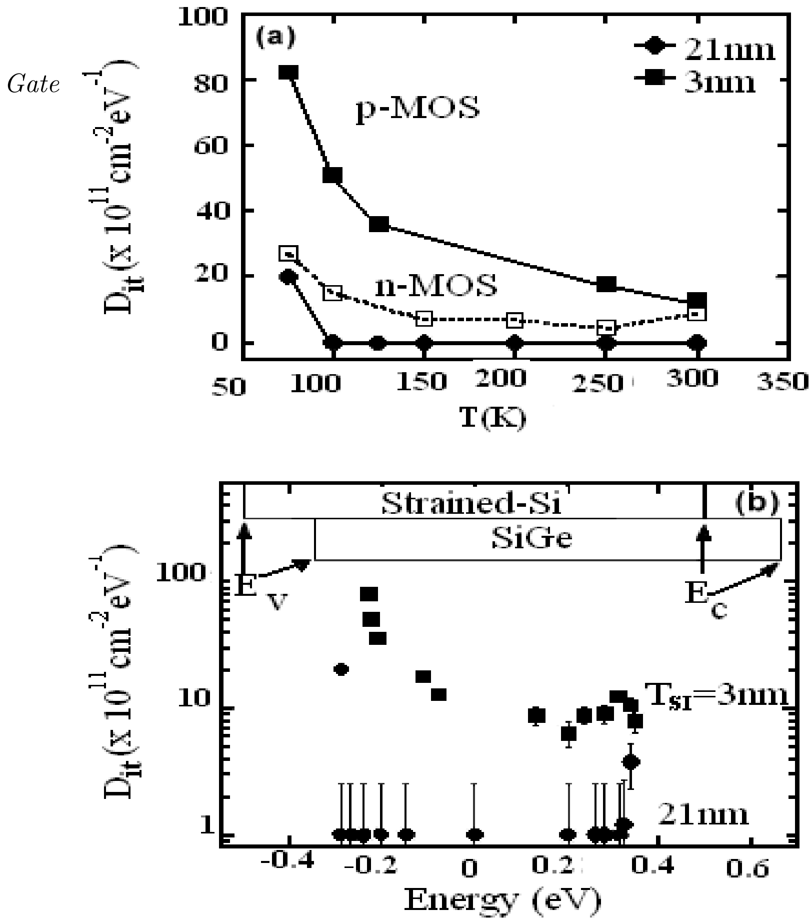



**FIGURE 5.24**
(a) Temperature dependence of $D_{it}$ obtained by the subthreshold swing of strained-SOIs at a $T_{Si}$ of 21 nm (circles) and 3 nm (squares). Open characteristics show the n-MOSFET data. At a $T_{Si}$ of 3 nm, $D_{it}$ of pMOS is much higher than that of n-MOSFETs, and $D_{it}$ rapidly increases when the temperature is lower than 150 K. (b) Interface state density profile considering the Fermi level shift at various temperatures. $D_{it}$ at the positive and the negative energy level is determined by the S-swings of n- and p-MOSFETs, respectively. At 3 nm (circles), $D_{it}$ near the valence band $E_v$ is much higher than $D_{it}$ near the conduction band $E_c$. After T. Mizuno et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(1114-1121). With permission.

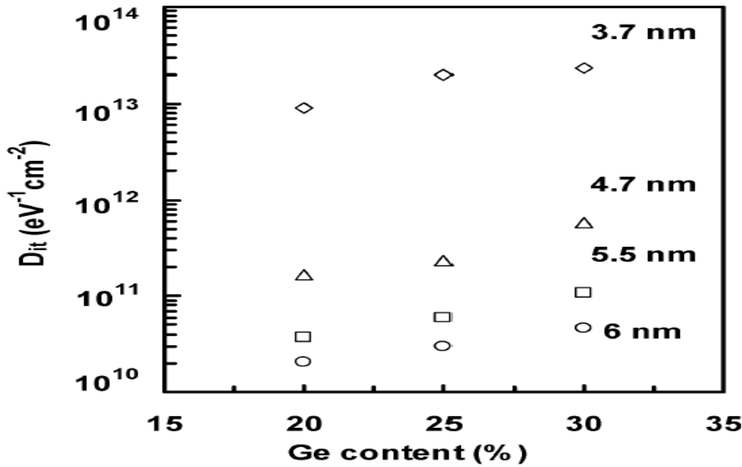



**FIGURE 5.25**
Variation in the interface state density as a function of Ge content in the virtual substrate. After G. K. Dalapati et al., *IEEE Trans. Electron Dev.*, Vol. 53, 2006(1142-1152). With permission.

In case of the buried strained-Si channel structures, the strained-Si/SiGe interface is considered to have fairly good quality. However, mobility lowering in thinner strained-Si has been explained by the increase in Coulomb scattering due to interface states in the strained-Si/gate-oxide, as follows. Fig. 5.23 shows the interface state density evaluated from the change in the subthreshold swing of strained-SOI MOSFETs. It was found that $D_{it}$ rapidly increases with decreasing $T_{Si}$, and is higher than $10^{12}$ cm$^{-2}$ eV$^{-1}$ in strained-Si layers that are thinner than 10 nm.

Dalapati et al. [22] have reported similar observations for strained-Si channel n-MOSFETs fabricated on virtual substrates with different Ge mole fractions. Fig. 5.25 shows the variation of $D_{it}$ with Ge content in the virtual substrate (VS) for different strained-Si channel thicknesses. The results indicate that the midgap $D_{it}$ of strained-Si/$SiO_2$ is higher than that of the bulk-Si/$SiO_2$, but has an acceptable value of about $3\times10^{10}$ eV$^{-1}$ cm$^{-2}$ for a device of channel thickness 6 nm and Ge content $< 30\%$. From Fig. 5.25, it may be noted that $D_{it}$ increases with Ge content in the VS for all strained-Si layer thicknesses. As the surface roughness of the strained-Si layer increases with Ge content in the VS, there is a strong relationship between $D_{it}$ and surface roughness of the strained-Si layer [56]. Thus, $D_{it}$ increases more rapidly as the strained-Si thickness is reduced (see Fig. 5.26). $D_{it}$ increases drastically below a strained-Si thickness of 4.7 nm, resulting in the Ge pile-up at the strained-Si/$SiO_2$ interface.

Fig. 5.24(a) shows the temperature dependence of $D_{it}$ values estimated

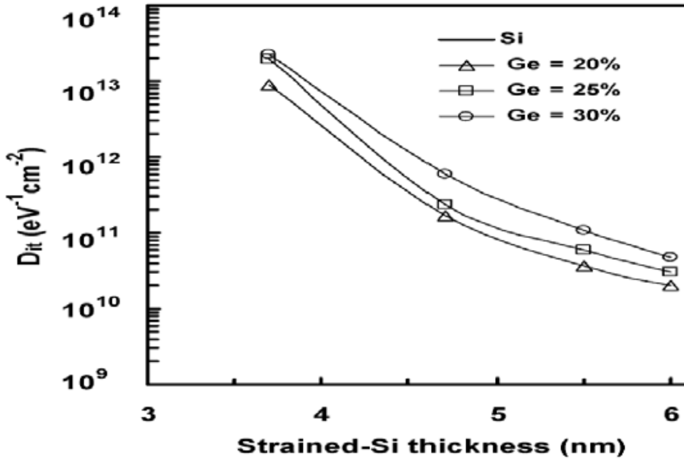



**FIGURE 5.26**
Variation in the interface state density as a function of strained-Si thickness of polySi/$SiO_2$ strained-Si MOS structure. After G. K. Dalapati et al., *IEEE Trans. Electron Dev.*, Vol. 53, 2006(1142-1152). With permission.

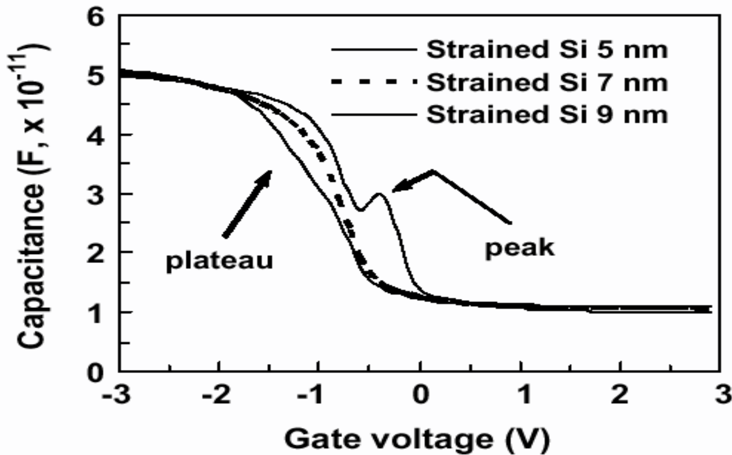



**FIGURE 5.27**
Capacitance-voltage (C-V) characteristics for the MOS structures with varying strained-Si cap layer thicknesses. The plateau in C-V curve indicates the existence of heterojunction while peak in the C-V curve around the depletion region indicates the presence of high-density traps. After S. Chattopadhyay et al., *Solid-State Electron.*, Vol. 48, 2004(1407-1416). With permission.

from subthreshold measurements for strained-SOI n-MOSFETs (open characteristics) and p-MOSFETs with 3 nm (squares) and 21 nm (circles) strained-Si layers. In a strained-Si thickness of 3 nm, $D_{it}$ of pMOS is much higher than that of n-MOS, and $D_{it}$ rapidly increases when temperature is lower than 150 K. Fig. 5.24(b) shows the energy distribution of $D_{it}$ of thick (circles) and thin strained-Si structures (squares), determined by the relationship between the $D_{it}$ value in Fig. 5.24(a) and the temperature-dependent positions of the Fermi level. The temperature dependence of the subthreshold swing at a given gate bias can be ascribed to the energy distribution of $D_{it}$ in the bandgap, which can be obtained by using the temperature dependence of the MOS surface potential. Fig. 5.24(b) shows that $D_{it}$ rapidly increases near the valance band edge, whereas it remains constant and relatively low in value near the conduction band edge.

Fig. 5.27 compares the high-frequency C-V characteristics for strained-Si MOS capacitors with 100×100 $\mu m^2$ area and varying strained-Si layer thicknesses of 9, 7, and 5 nm. Three different features are observed in the C-V curves with different strained-Si thicknesses. A plateau is observed in the C-V curve obtained from the capacitor with 9-nm strained-Si layer thickness, the curve for capacitor with 5-nm thickness shows a peak in the depletion region (- 0.4 V), and the curve for capacitor with 7-nm thickness shows no plateau or peak. The plateau in first case in the accumulation part is due to the confinement of holes and it indicates the presence of a quantum well at the heterojunction between the strained-Si and compressively strained-$Si_{0.7}Ge_{0.3}$. The peak in the second case is due to the presence of a large number of interface traps at the $SiO_2$/strained-Si interface. These observations confirm the higher Ge outdiffusion up to the Si/$SiO_2$ interface through the thin strained-Si layer.

## 5.7 High-k Gate Dielectrics on Strained-Si

Si-based metal oxide semiconductor field-effect transistors have been scaled close to their physical thickness limit. Aggressive scaling of gate oxide has reduced its thickness to a present-day value of about 1 nm, and any further thinning is prohibitive due to oxide reliability issues and the exponential increase of leakage current from direct tunneling. Sustaining Moore's law, however, depends crucially on the gate insulator scaling and consequently, replacing $SiO_2$ with high-k dielectric-metal gate stack is within sight. Silicon dioxide has been the gate dielectric in view of its amorphous structure, a bandgap of $> 9$ eV, band offsets relative to silicon greater than 1 eV. The gate oxide has to withstand the highest electric field. $SiO_2$ exhibits low $D_{it}$ ($\sim 5\times10^{10}$ states/$cm^2$-eV), low fixed charge density ($5\times10^{10}$ charges/$cm^2$), minimal low

**TABLE 5.3**
Dielectric constant, bandgap, electron and hole barrier heights for selected high-k gate dielectrics.

| | $SiO_2$ | $Si_3N_4$ | $HfO_2$ | $ZrO_2$ | $Al_2O_3$ | $TiO_2$ | $Ta_2O_5$ | $ZrSiO_4$ |
|---|---|---|---|---|---|---|---|---|
| Dielectric constant (k) | 3.9 | 7.0 | 25 | 20 | 9 | 80-100 | 25 | 10-12 |
| $E_g$ (eV) | 9 | 5.3 | 6 | 5.8 | 8.8 | 3.05 | 4.4 | 6 |
| $\Phi_c$ (eV) | 3.5 | 2.4 | 1.5 | 1.4 | 2.8 | 1.2 | 0.3 | 1.5 |
| $\Phi_v$ (eV) | 4.4 | 1.8 | 3.4 | 3.3 | 4.9 | – | 3.0 | 3.4 |

frequency capacitance-voltage (C-V) hysteresis and frequency dispersion, minimal dielectric charging and interface degradation or stress-induced leakage current (SILC) as a result of voltage stress or voltage bias-temperature measurements, sufficiently high mobility as well as exhibits sufficient reliability after years of high field stressing (> 10 years).

The gate dielectric thickness of $SiO_2$ has decreased from 50-100 nm in the 1970s to 1.2-1.5 nm equivalent oxide thickness (EOT) for state-of-the-art advanced MOSFETs, which is about 5 monolayers of silicon dioxide. The thinning of the gate dielectric, between 3 and 3.5 nm, necessary for the next generations of integrated devices, will give rise to unacceptably high gate leakage current arising from direct tunneling through the $SiO_2$ films. A typical leakage current density for 1.5 nm thick $SiO_2$ at 1 V is $\sim 1$ A/cm$^2$. There exists a physical limit for $SiO_2$ thickness (around 1.0-1.2 nm) as demonstrated at the atomic scale in a landmark experiment by Mueller et al. [57]. To further increase chip performance by using novel materials, the focus has been on high permittivity gate insulators and on metal gate electrodes. As $SiO_2$ will almost certainly be replaced by high-k gate dielectric in near future, the replacement of silicon as channel material by novel materials with improved transport properties is being viewed as a promising option to continue device scaling.

Since the introduction of Si MOSFETs in the 1960s, the semiconductor industry has seen an unprecedented rate of advancements in manufacturing and device performance. Much of this achievement is based on the scalability of the MOSFETs with silicon as the substrate and $SiO_2$ as the gate insulator. Now, after 40 years of decreasing size and increasing performance, traditional scaling is nearing fundamental limits at the 130 nm technology node [26]. There are two different avenues to achieve further enhancement in device performance: (a) by alternative device structures or (b) by looking for alternative material systems that exhibit higher mobility values. The main disadvantage of $SiO_2$ is its low permittivity (3.9). The permittivity of $SiO_2$ can be somewhat increased by mixing it with $SiN_y$. However, now a days, $SiO_xN_y$ is also approaching the limit where the gate oxide is about to be scaled so thin that the direct tunneling current through it would become unacceptably high.

At the 65-nm technology node, ultrathin $SiO_2$ gate dielectrics, of less than

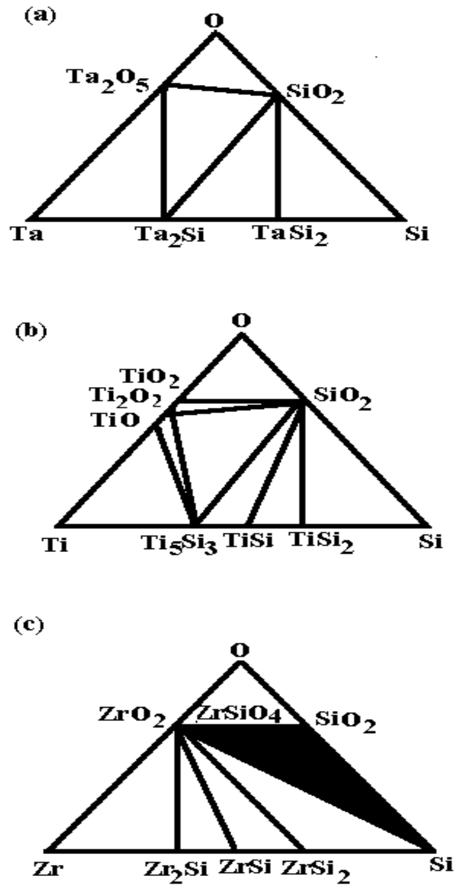



**FIGURE 5.28**
Phase diagrams show the oxide and silicate compounds made with Ta, Ti, and Zr. For the binary oxide to be stable in contact with silicon, a tie line must exist between the $MO_x$ (M = metal) and silicon. Unlike Ta- and Ti-based films, Zr compounds are stable in direct contact with silicon.

20 Å in thickness, will be needed. Such a thin oxide is prone to direct tunneling and exhibits a significant leakage current (typically 1 A/cm$^2$ at 2 V). In order to maintain a high drive current and also a minimum leakage current, a thicker film of high dielectric constant gate dielectric, with the same equivalent oxide thickness, may be used [58, 59]. At the present time, for the front-end processing, serious attempts are being made toward the development and integration of gate stack technologies which contain alternative high-k dielectrics. Electrical properties, such as dielectric constant, bandgap, electron, and hole barrier heights for selected high-k gate dielectrics are shown in Table 5.3. As the high-k gate dielectrics reduce gate leakage, and strained-Si enhances the performance of the MOSFET, it would be interesting to combine both for a solution toward the current limitations of Si CMOS downscaling [60, 61, 62].

Any potential high-k dielectric must be as thermodynamically stable on Si as $SiO_2$ or $SiO_xN_y$, which means that in further processing steps, following the gate dielectric film deposition, (a) chemical reaction in the bulk and at the high-k/Si interface must be prevented, (b) oxygen diffusion through the high-k film and oxygen reaction with the substrate-Si must be controlled, and (c) migration and eventual segregation of any element, especially at the electrode/dielectric and dielectric/Si interface, must be avoided [63].

Thermodynamic stability is the most critical factor in determining the adaptability of binary oxides as a gate dielectric because, otherwise, when in contact with Si, $SiO_2$ formation may be favorable and result in a degradation of the capacitance. Hubbard and Schlom [64] reported a thermodynamic study on most of the binary oxides to investigate their stability in contact with Si. Based on their work, the high-k material systems can be classified as a thermodynamically stable or unstable system.

The method employed by Hubbard and Schlom is to determine the M-Si-O phase diagrams, where M is an element of the periodic table of choice for silicon-compatible binary oxides [64]. The phase diagram contains silicon, the binary oxide ($MO_x$, M = metal), and the possible products of a reaction between silicon and the binary oxide. Although at high temperature, binary oxides have many possible reactions, considering the phase diagram of an M-Si-O system, one can identify several key reactions. An analysis of the free energies governing the relevant chemical reactions for ternary systems formed between tantalum (Ta), titanium (Ti) and zirconium (Zr) and silicon and oxygen gives the phase diagrams shown in Fig. 5.28. In the Ta-Si-O and Ti-Si-O ternary systems, as shown in Figs. 5.28(a) and (b), $Ta_2\,O_5$ and $TiO_2$, respectively, are not stable to $SiO_2$ formation when placed next to silicon [65].

In contrast to the Ta and Ti systems, the tie lines in the phase diagram for the Zr-Si-O system, indicate that the metal oxide $ZrO_2$ and the compound silicate $ZrSiO_4$ are both stable in direct contact with Si up to high temperatures. For the binary oxide to be stable in contact with silicon, a tie line must exist between the $MO_x$ and silicon. There are three possible types of M-Si-O phase diagrams. These are shown in Fig. 5.28 and referred to as "metal oxide dominant," "no phase dominant," and "$SiO_2$ dominant." Thermodynamic

data may be used to determine a particular M-Si-O system belongs to which of these three types of phase diagrams. According to the thermodynamic calculations, binary oxides shown in Table 5.3 are stable in contact with silicon [64].

Though from a thermodynamic point of view, some metal oxides are stable in contact with Si, this only implies that no reaction will occur if the pure metal oxide sits directly on top of Si under equilibrium conditions, which is definitely not the case during thin film deposition. The thin film deposition process, in general, involves nonequilibrium processes, where the kinetics of surface reactions will take over and determine the reaction products instead of equilibrium thermodynamic predictions. In fact, complicated deposition conditions and deposition ambient seldom make the deposition processes close to equilibrium, resulting in nonstoichiometric thin films [66].

Unlike the high temperature needed for the growth of $SiO_2$, the temperature used in most high-k dielectric deposition on Si can be made conveniently low, typically less than 600$^o$C. Different thin film deposition techniques have been employed for the deposition of high-k gate dielectrics. These techniques are usually divided into two main groups; physical vapor deposition (PVD) and chemical vapor deposition. In general, in all PVD methods no chemical exchange reactions should occur while in CVD the chemical reactions play a crucial role in the deposition process. Two typical PVD techniques are sputtering and evaporation, while metalorganic chemical vapor deposition (MOCVD) and atomic layer deposition (ALD) are two typical CVD techniques. To investigate the feasibility, both sputtering and CVD of $HfO_2$, $ZrO_2$, and $Ta_2O_5$ films have also been reported [58, 60, 67, 68].

It is also important to identify intrinsic limitations of the performance of alternative high-k gate dielectrics. The most important is specifically the introduction of an ultrathin interfacial $SiO_2$ layer between the Si substrate and the alternative dielectric. The interfacial layer can develop parasitically, during the direct deposition of Si oxynitride, binary oxides, and silicate on Si substrates. Alternatively, they can be introduced deliberately as ultrathin $SiO_x$ or nitrided $SiO_x$ buffer layers to reduce interfacial bonding defects, and/or to move fixed interfacial charge in the high-k away from the Si-dielectric interface so that it resides at an internal dielectric interface where its effect on threshold voltage and channel mobility is reduced.

However, to become a viable solution for $SiO_2$ replacement, alternative high-k gate dielectrics must exhibit low density of interface traps, excellent leakage performance and reliable transistor action, such as good mobility and high drive current. High-k gate dielectrics need careful evaluation for interface stability, thermal stability, process uniformity, and defect morphology/density. Deposition of high-k metal oxide dielectric layers on SiGe, SiGeC and strained-Si has been performed using microwave-plasma CVD of organometallic precursors [69, 70, 71, 72, 73, 74, 75, 76]. The oxides investigated so far, $ZrO_2$ and $HfO_2$, have dielectric constants in the range 20-30, and are thermodynamically stable on these epitaxial films. Growth of high-k metal oxide dielectric

layers on hydrophobic Ge (100) has also been performed using UV-ozone oxidation of vapor-deposited metal precursor films [77]. In fact, these oxides are sufficiently stable compared to $GeO_2$, that they appear to prevent formation of a low-k germanium oxide interface layer between the deposited dielectric and the Ge (100) surface. This is a key requirement for use of the dielectric in nanometer-scale transistors.

It now appears there is some consensus on the families of alternative gate materials that will likely to be used in future CMOS technologies. $HfO_2$ and $ZrO_2$ have particular potential, due to their compatibility with both polySi and TaN gates. For hafnium and zirconium, major focus is now on the pure metal oxides and their silicates and aluminates. Although $Al_2O_3$ has a high crystallization temperature, a large bandgap and band offset, a low leakage current, its k value is too low for the desired scaling of the equivalent oxide thickness.

### 5.7.1 Microwave Plasma Deposition

Among the high-k dielectrics that are currently being considered as promising candidates are $Ta_2O_5$, $ZrO_2$ and $TiO_2$, and $HfO_2$. Mixed oxides and silicates are also gaining interest as alternative high-k gate dielectrics. In the following, we present the results of some study on the deposition and electrical characterization of several high-k gate dielectrics on strained-Si on relaxed-SiGe buffer. Emphasis will be given to the current conduction mechanisms in these deposited gate dielectrics.

Prior to deposition, the strained-Si layers (on relaxed-SiGe) grown using gas source molecular beam epitaxy, were chemically etched in $H_2O_2$:$H_2SO_4$ solution followed by a dip in 1% HF solution to remove the native oxide layer. All the deposition was carried out in a microwave plasma chemical vapor deposition system [78] using appropriate metallorganic precursor. For the deposition of $SiO_xN_y/Ta_2O_5$ gate stack, $SiO_xN_y$ layer was grown in a microwave plasma (pressure 400 mTorr and temperature 150$^o$C) by exposing the strained-Si substrate to nitric oxide. Then the tantalum pentaethoxide [$Ta(OC_2H_5)_5$], vaporized from a bubbler kept at 150$^o$C, was carried to the quartz process/deposition chamber of the microwave (700 W, 2.45 GHz) cavity discharge system via a heated gas line to prevent condensation. The process chamber was maintained at a pressure of 400 mTorr during deposition. During the $Ta(OC_2H_5)_5$ decomposition, a complex set of chemical reactions takes place, starting from the precursor solution (during film formation) until an oxide network of metal-oxygen-metal bond is produced. In the temperature range of 150-275$^o$C, a self-controlled growth mode was achieved as physisorption is suppressed. During the CVD process, oxide molecules are also formed from the pyrolysis of an alkoxide molecule, which basically is a thermal decomposition of the metallorganic compound into its component elements as shown in the following reaction: [79]

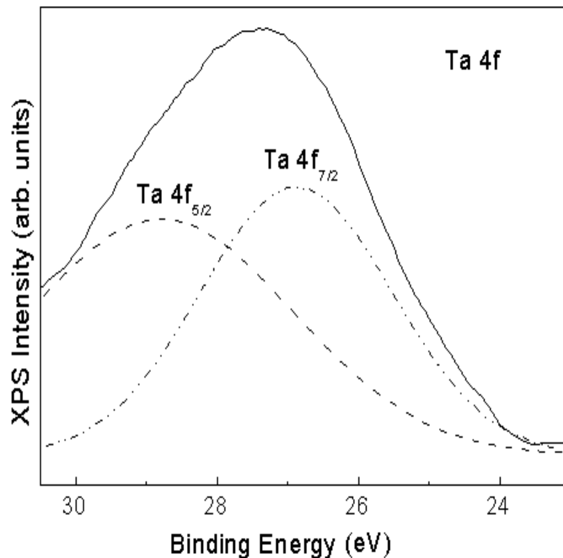
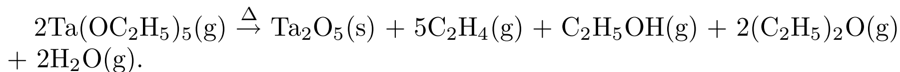



**FIGURE 5.29**
X-ray photoelectron spectroscopy of the Ta 4f peaks in $Ta_2O_5$ films.



$$2Ta(OC_2H_5)_5(g) \overset{\Delta}{\rightarrow} Ta_2O_5(s) + 5C_2H_4(g) + C_2H_5OH(g) + 2(C_2H_5)_2O(g) + 2H_2O(g).$$

For the deposition of $TiO_2$ films, organometallic titanium tetrakis isopropoxide (TTIP) was used as a precursor and bubbler temperature was maintained at 45$^o$C. Zirconium tetratert butoxide $[Zr(OC(CH_3)_3)_4]$ was used as a source material for the deposition of $ZrO_2$ films and was introduced in the deposition chamber of the microwave cavity discharge system independently from a bubbler kept at room temperature.

The film thicknesses were measured using a single wavelength (6328 Å) ellipsometer (model Gaertner L-117). The layer thicknesses were found to be typically in the range of 120-140 Å for a typical deposition time of 90 s. For the electrical measurements, metal insulator semiconductor (MIS) test capacitors were fabricated using the deposited dielectric layers with evaporated circular aluminum contacts (area: $1.96\times10^{-3}$ cm$^2$) through a shadow mask. The capacitance-voltage and current-voltage characteristics (including under constant current stressing) were studied using the HP-4061A semiconductor test system and an HP-4145B DC parameter analyzer, respectively.

## 5.7.2 Chemical Analysis

The surface chemical states of the as-deposited films were analyzed by x-ray photoelectron spectroscopy. Nonmonochromatized Mg $K_\alpha$ (h$\gamma$ = 1253.6 eV)

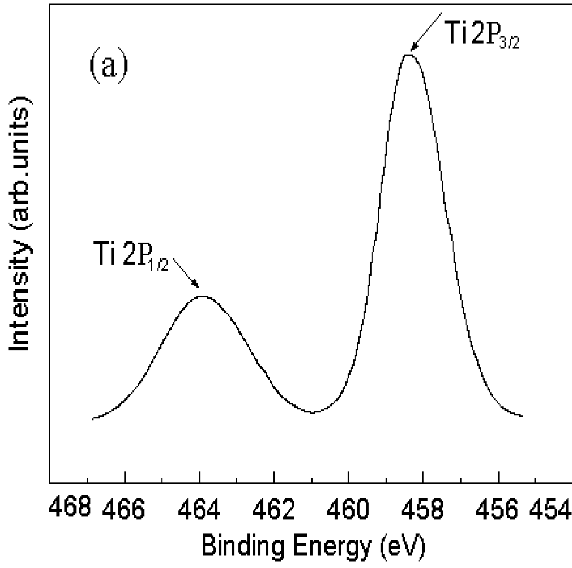



**FIGURE 5.30**
XPS spectra of Ti 2P for the as-deposited $TiO_2$ films on strained-Si layers.

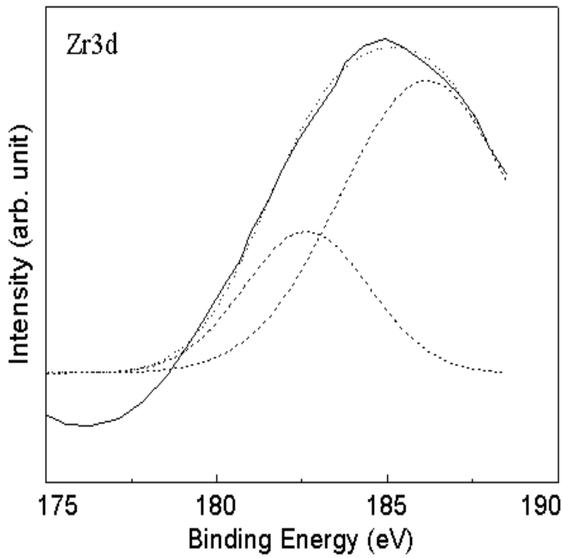



**FIGURE 5.31**
Deconvoluted Zr 3d spectra of plasma-deposited $ZrO_2$ films on strained-Si.

**TABLE 5.4**
Electrical properties of several microwave plasma deposited high-k gate dielectric films on strained-Si/SiGe heterolayers.

| Substrate | Oxide type | Process used | $Q_f/q$ ($cm^{-2}$) | $D_{it}$ ($cm^{-2}eV^{-1}$) ) | Ref. |
|---|---|---|---|---|---|
| strained-Si | $SiO_xN_y$ /$Ta_2O_5$ | MWCVD | $4.63\times10^{11}$ | $2.3\times10^{11}$ | [61] |
| strained-Si | $TiO_2$ | MWCVD | $1.38\times10^{11}$ | $1.18\times10^{12}$ | [85] |
| strained-Si | $ZrO_2$ | MWCVD | $1.15\times10^{11}$ | $2.24\times10^{12}$ | [86] |
| strained-Si | $ZrO_2$ | MWCVD | $1.2\times10^{11}$ | $3\times10^{11}$ | [35] |
| SiGe (x = 26%) | $HfO_2$ | rf sputtering | $1.4\times10^{12}$ | $5.5\times10^{11}$ | [87] |
| SiGe (x = 18%) | $HfO_2$ | rf sputtering | - | - | [68] |

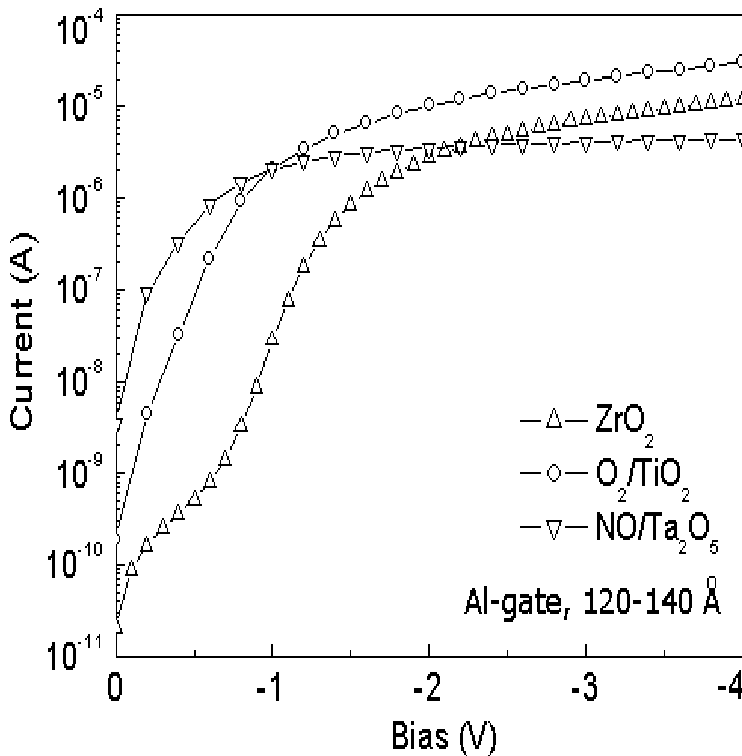



**FIGURE 5.32**
I-V characteristics of several microwave plasma-deposited high-k gate dielectrics. The leakage current is reduced by more than three orders of magnitude compared to $SiO_2$.

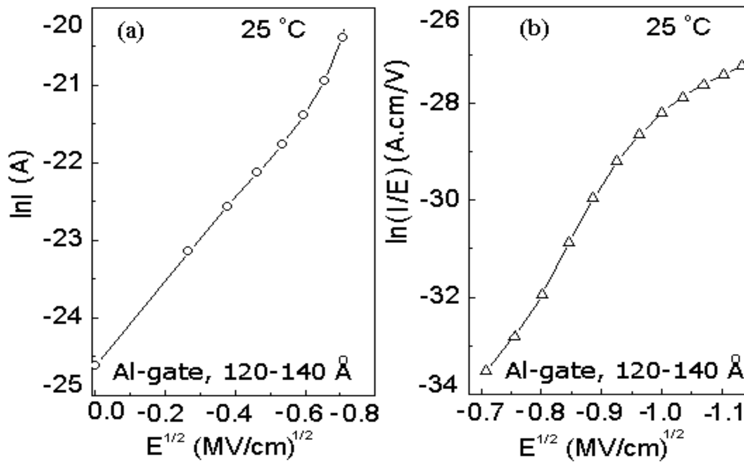



**FIGURE 5.33**
Current conduction mechanisms in $ZrO_2$ films: (a) Schottky emission and (b) Poole-Frenkel.

radiation was used at an angle $30^o$ between the analyzer axis and normal to the sample to investigate the chemical structure of the films from the shift of the core-level binding energy. For the $Ta_2O_5$ films, the Ta 4f XPS peak (see Fig. 5.29) is deconvoluted into two subpeaks. The two peaks are located at 28.7 eV and 26.8 eV with energy separation of 1.9 eV for Ta $4f_{5/2}$ and Ta $4f_{7/2}$ state, respectively. The Ta 4f peak shift of 0.6 eV toward higher binding energy is attributed to the increasing oxygen content of the tantalum oxide and the Ta $4f_{7/2}$ peak position at 26.8 eV is typical of $Ta^{+5}$ chemical state in $Ta_2O_5$ [80].

For the $TiO_2$ films, to investigate the bonding chemistry, the Ti 2P spectra were deconvoluted rather than the oxygen spectra because the peak position of O 1S in $TiO_2$ is the same as that of in TiO and $Ti_2O_3$. Fig. 5.30 shows the Ti 2P XPS spectra of the as-deposited $TiO_2$ films. The binding energy of Ti $2P_{3/2}$ and Ti $2P_{1/2}$ were observed at 458.3 and 463.9 eV, respectively, with a separation of 5.6 eV between the peaks, which is a typical characteristic of the $Ti^{4+}$ in $TiO_2$. Binding energy of O 1S for $SiO_2$ appears at 533.5 eV and that for $TiO_2$ appears at 530 eV.

For the $ZrO_2$ films, the XPS spectrum shows that the plasma-deposited $ZrO_2$ film has a typical $ZrO_2$ chemical bonding structure (see Fig. 5.31). The Zr 3d spectra show two peaks (deconvoluted) at 183.1 and 185.8 eV binding energy and are both characteristic of fully oxidized Zr. Similar results have also been reported by Miyazaki et al. [81] for $ZrO_2$ films deposited directly on Si. Guittet et al. [82] reported that the Zr 3d state binding energy is more than 0.5 eV larger in $ZrSiO_4$ than in $ZrO_2$, but there is no such peak for $ZrSiO_4$

in the film. The evidence of nonexistence of Zr-Si bonding suggests that Zr is oxidized before it could react with Si to form Zr-Si bonds in ultrathin films. The equivalent oxide thickness of the gate dielectric films were calculated using the following expression [83]

$$EOT = \epsilon_{SiO_2} T_{int}/\epsilon_{int} + \epsilon_{SiO_2} T_{hk}/\epsilon_{hk} \tag{5.3}$$

where $T_{int}$ and $T_{hk}$ are the thicknesses of the interfacial layer and high-k film, and $\epsilon_{SiO_2}$, $\epsilon_{int}$ and $\epsilon_{hk}$ are the dielectric permittivities of the $SiO_2$, the interfacial layer and the high-k film, respectively.

Fig. 5.32 shows the typical I-V characteristics of the MOS capacitors using various high-k gate dielectrics. Compared to $SiO_2$, the leakage current is reduced by more than three orders of magnitude for all the high-k gate dielectrics. While thicker high-k layers can be used to prevent direct tunneling between the gate and substrate, thermionic emission is a serious concern, particularly if the band alignment is not favorable. In general, with increasing atomic number of the metal, the atomic (or ionic) radius increases and so the cohesive force decreases, leading to a high dielectric constant but a narrow energy gap [84]. As expected, the stress-induced leakage current (SILC) for a constant current stressing of 300 s and 500 s, indicates a higher leakage current at a low field due to stressing. The breakdown field for some of these MOS capacitors was found to be between 2 and 5 MV/cm. The hysteresis due to change of voltage sweep direction was found to be below 20 mV for all the dielectric films. Table 5.4 summarizes some of the important electrical properties for the microwave plasma-deposited high-k dielectrics.

### 5.7.3 Conduction Mechanism

To explain the current transport in ultrathin oxides, the Schottky emission (S-E), Poole-Frenkel (P-F) emission, Fowler-Nordheim (F-N) tunneling, direct tunneling, and space charge limited current mechanisms are commonly employed. The Poole-Frenkel effect has been used to explain the conduction in thin dielectric films with reasonable accuracy. In addition, recent investigations into thin films with high dielectric constants, have shown that the current flow in these materials is bulk-limited by the P-F effect. The F-N tunneling is usually considered to be the dominant conduction mechanism in insulator of a MIS capacitor at a very high field on the order of 7-8 MV/cm, whereas the P-F conduction is observed at a low field.

In the Poole-Frenkel conduction, it is known that the field enhances the thermal excitations and lowers the Coulomb potential for traps in the oxide to emit carriers to the conduction band of the oxide. The dominance of this trap-assisted conduction mechanism comes from the fact that the dielectric films are much thicker than the interfacial layer of $SiO_2$, where the carrier may tunnel through directly.

The current, $I_{SE}$ due to S-E is expressed as follows: [88]

$$I_{SE} = AT^2 exp\frac{-(q\Phi - \beta\sqrt{E})}{kT} \tag{5.4}$$

$$\beta_{SE} = \sqrt{\frac{q^3}{4\pi\epsilon_o\epsilon_r}} \tag{5.5}$$

The gradient of the S-E, M, is proportional to $\beta_{SE}$, and from Eqn. 5.5, one can express M as

$$M = \frac{\beta_{SE}}{\xi kT} \tag{5.6}$$

where A is a constant, $\Phi$ the Schottky barrier height, $\epsilon_r$ is the dielectric constant of the insulator, $\epsilon_o$ is the permittivity of the free space and E is the electric field. The S-E mechanism gets modified to the P-F mechanism if the conduction is governed by carriers that are thermally emitted from trapped centers under a strong electric field.

The Poole-Frenkel effect is often used to explain electronic conduction in thin dielectric films. A first-order model [89] for the current due to the P-F effect has been used to explain the I-V characteristics. For the electron to be emitted from the trap, and be available for conduction, it must gain sufficient energy to surmount the peak of the potential energy barrier. The general expression for the P-F current ($I_{PF}$) is given by: [88]

$$I_{PF} = CEexp[\frac{-(q\Phi - \beta_{PF}\sqrt{E})}{\xi kT}] \tag{5.7}$$

where C and k are the proportionality constant and Boltzmann constant, respectively. q$\Phi$ is the ionization potential in eV, which is the amount of energy required for the trap electron to escape the influence of the positive nucleus of the trapping center when no field is applied. $\beta_{PF}\sqrt{E}$ is the amount by which the trap barrier height is reduced by the E. The factor $\xi$ in the denominator of the exponential may vary between 1 and 2, depending on the amount of acceptor concentration.

The Poole-Frenkel constant, $\beta_{PF}$ is given by

$$\beta_{PF} = \sqrt{\frac{q^3}{\pi\epsilon_o\epsilon_r}} \tag{5.8}$$

$$ln[\frac{I}{E}] = \frac{\beta_{PF}}{\xi kT} \cdot \sqrt{E} + [lnC - \frac{q\Phi}{\xi kT}] \tag{5.9}$$

which is also referred to as the P-F plot, and a region of linearity in an experimental plot of ln[I/E] vs. $\sqrt{E}$ is taken as an indication of P-F conduction. The gradient of the P-F, $M_1$, is proportional to $\beta_{PF}$, and from Eqn. 5.9, one can express $M_1$ as

$$M_1 = \frac{\beta_{PF}}{\xi kT} \tag{5.10}$$

In the following, the leakage current conduction mechanism in microwave plasma-deposited ultrathin high-k gate dielectrics films at room temperature is discussed. The logarithm of the current for the $ZrO_2$ films is plotted against the square root of electric field and is shown in Fig. 5.33(a). A straight line, obtained at low electric fields ($\leq$ 0.7 MV/cm)$^{1/2}$, implies the Schottky emission. At a higher electric field, the conduction is governed, however, by a different mechanism. In Fig. 5.33(b), the logarithmic current divided by the electric field is plotted as a function of the square root of the electric field [ln(I/E) vs. $E^{1/2}$]. A straight line is obtained also at a higher electric field ($\geq$ 0.7 MV/cm)$^{1/2}$, suggesting that the leakage current at a high electric field is due to the P-F conduction mechanism. From Fig. 5.33(a), the value of $M_1$ is found to be $5.4\times10^{-3}$.

In order to study the temperature dependence of interface states in detail for the $SiO_xN_y/Ta_2O_5$ stack, I-V characteristics were measured at several temperatures in the range of 25-160$^o$C and is shown in Fig. 5.34. To determine the dominant leakage current mechanism, a plot of ln(J/E) vs. $E^{1/2}$ (P-F mechanism) has been used (Fig. 5.35). $SiO_xN_y/Ta_2O_5$ gate stack clearly shows a linear relationship and the conduction mechanism is dominated by the P-F mechanism at a low field and high temperature.

### 5.7.4 Reliability Issues

High-k gate dielectrics are known to be "trap-rich" materials. Relatively thick high-k dielectrics provide researchers a unique opportunity to study the dielectric degradation process under extremely high electric field under different conditions such as Fowler-Nordheim and direct tunneling current injection [90, 91, 92]. Efforts are being made to improve high-k dielectrics characteristics. For example, $NH_3$, forming gas anneals and $O_3$ surface treatment have been shown to reduce the interface states [93, 94]. The origin of traps in high-k dielectrics, however, still remains a question. In order to evaluate reliability, it is necessary to pinpoint the factors that influence the breakdown of high-k dielectrics. The parameters that are experimentally measured to evaluate the reliability of high-k gate materials from MOS devices must be carefully collected by statistical methods.

The dielectric reliability of the high-k gates are evaluated by measuring the time-to-breakdown during constant voltage or constant current stressing [95]. It is important to study the issues such as soft and hard breakdown. As the physical and chemical nature of the interface layer plays an important role, defect generation rates, polarity dependence due to asymmetric band structure, charging effect by pre-existing traps, thickness dependence, area scaling, and AC stressing studies are also necessary [96]. Most gate dielectric reliability testing makes use of DC stressing because of simplicity. However, in an actual circuit, gate dielectrics are, in general, under dynamic stressing which, however, leads to higher time-to-breakdown than static DC stressing due to reduction of charge trapping in the dielectrics.

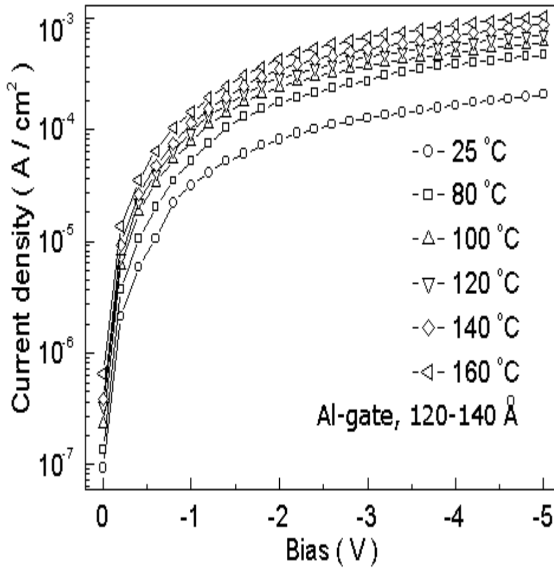



**FIGURE 5.34**
I-V characteristics of $SiO_xN_y/Ta_2O_5$ gate stack at different temperatures. The I-V data shows Poole-Frenkel emission over the whole range until the breakdown voltage.

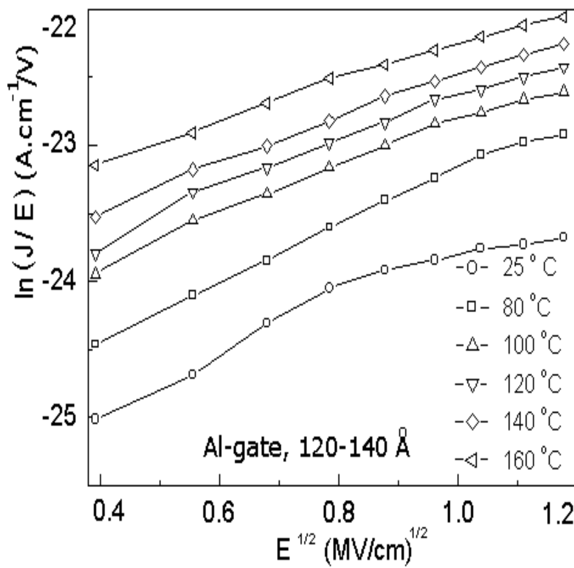



**FIGURE 5.35**
Current conduction mechanism in $SiO_xN_y/Ta_2O_5$ gate stack: Poole-Frenkel.

Till now, the reliability studies on the high-k gate dielectrics have raised many questions. So far the $SiO_2$-dominated breakdown mechanisms have been extended to high-k. It is still not known how the microscopic structure of the high-k layer influence degradation and breakdown. And finally, to what extent is the knowledge acquired on $SiO_2$ reliability transferable to high-k gate dielectrics.

## 5.8 Gate Dielectrics on Ge

Since Ge inversion layer mobility is the key metric in comparison to traditional Si transistor technology, the integration of a reliable gate stack is very important in exploiting the benefits of Ge transistor technology. Germanium oxynitride [97], aluminum oxide, hafnium oxide, and zirconium oxide [98] high-k dielectrics have been implemented with Ge channel transistors with varying degrees of success. However, as with the integration of high-k gate stacks with Si, no completely satisfactory solution has been demonstrated with Ge channels. This in part has led to less than anticipated mobility enhancements for holes and typical electron mobilities that are inferior to that of Si. For methods that use epitaxial Ge layers, thin layers of Si, which preserve the gate oxide integrity, have been used to cap the Ge channel layers. Relaxed and biaxially strained embodiments of Ge transistors using Si capping layers have markedly better mobility characteristics (2-10$\times$) than Si control transistors [99]. Some questions about properties of Ge in such applications require further analysis. Ge has a smaller bandgap than Si, which can cause junction leakage problems. In addition, the saturated high field drift velocity of charge carriers in Ge is somewhat lower than that of Si. Further experiments and simulations are needed to fully establish the benefits and trade-offs of modern Ge electronics. As the benefits are established, the optimum substrate and transistor embodiments will become clearer.

## 5.9 Summary

In this chapter, we have discussed the strain state, film, and surface morphology of strained-Si layers on SiGe virtual substrates grown by GSMBE and UHVCVD techniques. It is shown that compared to solid-source molecular beam epitaxy or chemical vapor deposition, the surface roughness increases significantly in the GSMBE-grown SiGe virtual substrates as the Ge concentration and/or the growth temperature exceeds 600$^o$C. We have dis-

cussed the kinetics and mechanisms of thermal and rapid thermal oxidation of strained-Si/SiGe heterolayers. Beyond 800$^o$C, Ge is found to diffuse into strained-Si layer. It has been shown that a large-scale cross-hatching, inherent in a relaxed-SiGe layer, leads to an increase in strained-Si/$SiO_2$ interface roughness. The surface morphology of the graded buffer needs to be improved. Modification of the epitaxial growth methods to minimize the undulating SiGe morphology and chemical mechanical polishing are expected to improve the strained-Si surface.

The oxidation of strained-Si/SiGe heterolayers by microwave plasma offers several advantages over conventional thermal oxidation. Movement of Ge atoms into strained-Si is found to be minimum for rapid thermal annealing at 800$^o$C for 30 min or at 900$^o$C for 3 min. Significant atomic scale diffusion of Ge into the strained-Si channel will result otherwise, leading to a marked reduction in the electron mobility. However, the introduction of SiGe-free strained-Si films on insulators is promising for a wider process temperature window.

Electrical properties of the microwave plasma-deposited high-k gate dielectrics have been presented, as well as microstructures of the oxide/substrate, and strained-Si/SiGe/Si substrate interfaces. It is shown that the high-k dielectrics on strained-Si layers can make it possible to avoid, in addition to the large leakage current, the strain relaxation and dislocation propagation. It is important to note that, to suppress these problems in strained layers, all processing temperatures and time should be carefully optimized according to the device needs, such as the oxide thickness, degree of strain, and the thickness of the strained layer.

Interfacial and electrical properties of the deposited high-k gate dielectrics such as $Ta_2O_5$, $TiO_2$, and $ZrO_2$ films were characterized using capacitance-voltage, current-voltage and conductance-voltage techniques. The C-V characteristics exhibit a frequency dependence due to the presence of interface traps. The leakage current conduction has been found to be dominated, in general, by the Schottky emission at a low electric field, whereas Poole-Frenkel emission takes over at a higher electric field.

## References

[1] C. K. Maiti, N. B. Chakrabarti, and S. K. Ray, *Strained Silicon Heterostructures: Materials and Devices.* The IEE, UK, 2001.

[2] S. F. Nelson, K. Ismail, T. N. Jackson, J. J. Nocera, and J. O. Chu, "Systematics of electron mobility in Si/SiGe heterostructures," *Appl. Phys. Lett.*, vol. 63, pp. 794–796, 1993.

[3] K. Ismail, B. S. Mayerson, S. Rishton, J. Chu, S. Nelson, and J. Nocera, "High-transconductance n-type Si/SiGe modulation-doped field-effect transistors," *IEEE Electron Dev. Lett.*, vol. 13, pp. 229–231, 1992.

[4] U. Konig, A. J. Boers, F. Schaffler, and E. Kasper, "Enhancement mode n-channel Si/SiGe MODFET with high intrinsic transconductance," *Electron. Lett.*, vol. 28, pp. 160–162, 1992.

[5] D. K. Nayak, K. Goto, A. Yutani, J. Murota, and Y. Shiraki, "High-mobility strained-Si PMOSFETs," *IEEE Trans. Electron Dev.*, vol. 43, pp. 1709–1715, 1996.

[6] J. Welser, J. L. Hoyt, and J. F. Gibbons, "Growth and processing of relaxed-$Si_{1-x}Ge_x$/strained-Si structures for metal oxide semiconductor applications," *Jap. J. Appl. Phys.*, vol. 33, pp. 2419–2422, 1994.

[7] J. Welser, J. L. Hoyt, and J. F. Gibbons, "Growth and processing of relaxed-$Si_{1-x}Ge_x$/strained-Si structures for MOS applications," *Proc. Int'l Conf. on SSDM*, pp. 377–379, 1993.

[8] K. Rim, J. Welser, J. L. Hoyt, and J. F. Gibbons, "Enhanced hole mobilities in surface-channel strained-Si p-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 517–520, 1995.

[9] S. H. Olsen, A. G. O'Neill, L. S. Driscoll, K. S. K. Kwa, S. Chattopadhyay, A. M. Waite, Y. T. Tang, A. G. R. Evans, D. J. Norris, A. G. Cullis, D. J. Paul, and D. J. Robbins, "High-performance nMOSFETs using a novel strained Si/SiGe CMOS architecture," *IEEE Trans. Electron Dev.*, vol. 50, pp. 1961–1969, 2003.

[10] L. K. Bera, S. K. Ray, D. K. Nayak, N. Usami, Y. Shiraki, and C. K. Maiti, "Electrical properties of oxides grown on strained-Si using microwave $N_2O$ plasma," *Appl. Phys. Lett.*, vol. 70, pp. 66–68, 1997.

[11] L. K. Bera, *Studies on applications of strained-Si for heterostructure field effect transistors (HFETs).* Ph.D. thesis, IIT-Kharagpur, 1998.

[12] N. Sugii, "Thermal stability of the strained-Si/$Si_{0.7}Ge_{0.3}$ heterostructure," *J. Appl. Phys.*, vol. 89, pp. 6459–6463, 2001.

[13] N. R. Zangenberg, J. L. Hansen, J. Fage-Pedersen, and A. N. Larsen, "Ge self-diffusion in epitaxial $Si_{1-x}Ge_x$ layers," *Phys. Rev. Lett.*, vol. 87, pp. 125901–1–125901–4, 2001.

[14] D. K. Nayak, K. Kamjoo, J. S. Park, J. C. S. Woo, and K. L. Wang, "Rapid thermal processing of strained GeSi layers," *IEEE Trans. Electron Dev.*, vol. 39, pp. 56–63, 1992.

[15] D. K. Nayak, K. Kamjoo, J. S. Park, J. C. S. Woo, and K. L. Wang, "Wet oxidation of GeSi strained layers by rapid thermal processing," *Appl. Phys. Lett.*, vol. 57, pp. 369–371, 1990.

[16] F. K. LeGoues, R. Rosenberg, T. Nguyen, F. Himpsel, and B. S. Meyerson, “Oxidation studies of SiGe,” *J. Appl. Phys.*, vol. 65, pp. 1724–1728, 1989.

[17] D. Fathy, O. W. Holland, and C. W. White, “Formation of epitaxial layers of Ge on Si substrates by Ge implantation and oxidation,” *Appl. Phys. Lett.*, vol. 51, pp. 1337–1339, 1987.

[18] E. C. Frey, N. Yu, B. Patnaik, N. R. Parikh, M. L. Swanson, and W. K. Chu, “Transition between Ge segregation and trapping during high-pressure oxidation of $Ge_xSi_{1-x}$/Si,” *J. Appl. Phys.*, vol. 74, pp. 4750–4755, 1993.

[19] D. C. Paine, C. Caragianis, and A. F. Schwartzman, “Oxidation of $Si_{l-x}Ge_x$ alloys at atmospheric and elevated pressure,” *J. Appl. Phys.*, vol. 70, pp. 5076–5084, 1991.

[20] P.-E. Hellberg, S.-L. Zhang, F. M. d’Heurle, and C. S. Petersson, “Oxidation of silicon-germanium alloys. I. An experimental study,” *J. Appl. Phys.*, vol. 82, pp. 5773–5778, 1997.

[21] C. Tetelin, X. Wallart, L. Vescan, and J. P. Nys, “Plasma assisted oxidation of SiGe layers at $500^oC$: interface characterization,” *Appl. Surf. Sci.*, vol. 104/105, pp. 385–391, 1996.

[22] G. K. Dalapati, S. Chattopadhyay, K. S. K. Kwa, S. H. Olsen, Y. L. Tsang, R. Agaiby, A. G. O’Neill, P. Dobrosz, and S. J. Bull, “Impact of strained-Si thickness and Ge outdiffusion on gate oxide quality for strained-Si surface channel n-MOSFETs,” *IEEE Trans. Electron Dev.*, vol. 53, pp. 1142–1152, 2006.

[23] T. Mizuno, N. Sugiyama, T. Tezuka, T. Numata, and S. Takagi, “High-performance strained-SOI CMOS devices using thin film SiGe-on-insulator technology,” *IEEE Trans. Electron Dev.*, vol. 50, pp. 988–994, 2003.

[24] F. K. Legoues, R. Rosenberg, and B. S. Meyerson, “Kinetics and mechanism of oxidation of SiGe: dry versus wet oxidation,” *Appl. Phys. Lett.*, vol. 54, pp. 644–646, 1989.

[25] F. K. LeGoues, B. S. Meyerson, and J. F. Morar, “Anomalous strain relaxation in SiGe films and superlattices,” *Phys. Rev. B*, vol. 66, pp. 2903–2906, 1991.

[26] Semiconductor Industry Association, “International Technology Roadmap for Semiconductors.” San Jose, Calif., 2005.

[27] S. K. Samanta, S. Chatterjee, S. Maikap, L. K. Bera, H. D. Banerjee, and C. K. Maiti, “Interface properties and reliability of ultrathin oxynitride films grown on strained-$Si_{1-x}Ge_x$ substrates,” *J. Appl. Phys.*, vol. 93, pp. 2464–2471, 2003.

[28] T. Hori and H. Iwasaki, "Improved hot-carrier immunity in submicrometer MOSFETs with reoxidized nitrided oxides prepared by rapid thermal processing," *IEEE Electron Dev. Lett.*, vol. 10, pp. 64–67, 1989.

[29] T. Ito, T. Nakamura, and H. Ishikawa, "Advantages of thermal nitridation and nitroxide gate films in VLSI process," *IEEE Trans. Electron Dev.*, vol. ED-29, pp. 498–502, 1982.

[30] M. M. Moslehi and K. C. Saraswat, "Thermal nitridation of Si and $SiO_2$ for VLSI," *IEEE Trans. Electron Dev.*, vol. ED-32, pp. 106–123, 1985.

[31] B. S. Doyle and G. J. Dunn, "Dynamic hot-carrier stressing of reoxidized nitrided oxide," *IEEE Electron Dev. Lett.*, vol. 12, pp. 63–65, 1991.

[32] T. Hori, H. Iwasaki, and K. Tsuji, "Electrical and physical characteristics of ultrathin reoxidized nitrided oxides prepared by rapid thermal processing," *IEEE Trans. Electron Dev.*, vol. 36, pp. 340–350, 1989.

[33] H. Hwang, W. Ting, B. Maiti, D. L. Kwong, and J. Lee, "Electrical characteristics of ultrathin oxynitride gate dielectric prepared by rapid thermal oxidation of Si in $N_2O$," *Appl. Phys. Lett.*, vol. 57, pp. 1010–1011, 1990.

[34] L. K. Bera, S. K. Ray, M. Mukhopadhyay, D. K. Nayak, N. Usami, Y. Shiraki, and C. K. Maiti, "Electrical properties of $N_2O/NH_3$ plasma grown oxynitride on strained-Si," *IEEE Electron Dev. Lett.*, vol. 19, pp. 273–275, 1998.

[35] C. S. Tan, W. K. Choi, L. K. Bera, K. L. Pey, D. A. Antoniadis, E. A. Fitzgerald, M. T. Currie, and C. K. Maiti, "$N_2O$ oxidation of strained-Si/relaxed-SiGe heterostructure grown by UHVCVD," *Solid-State Electron.*, vol. 45, pp. 1945–1949, 2001.

[36] H. Klauk, T. N. Jackson, S. F. Nelson, and J. O. Chu, "Thermal stability of undoped strained-Si channel SiGe heterostructures," *Appl. Phys. Lett.*, vol. 68, pp. 1975–1977, 1996.

[37] T. S. Drake, C. NyChleirigh, M. L. Lee, A. J. Pitera, E. A. Fitzgerald, D. A. Antoniadis, D. H. Anjum, J. Li, R. Hull, N. Klymko, and J. L. Hoyt, "Effect of rapid thermal annealing on strain in ultrathin strained silicon on insulator layers," *Appl. Phys. Lett.*, pp. 875–877, 2003.

[38] C. Tetelin, X. Wallart, J. P. Nys, L. Vescan, and D. J. Gravesteijn, "Kinetics and mechanism of low temperature atomic oxygen-assisted oxidation of SiGe layers," *J. Appl. Phys.*, vol. 83, pp. 2842–2846, 1998.

[39] P. W. Li, H. K. Liou, E. S. Yang, S. S. Iyer, T. P. Smith, and Z. Lu, "Formation of stoichiometric SiGe oxide by electron cyclotron resonance plasma," *Appl. Phys. Lett.*, vol. 60, pp. 3265–3267, 1992.

[40] M. Mukhopadhyay, S. K. Ray, C. K. Maiti, D. K. Nayak, and Y. Shiraki, "Electrical properties of oxides grown on strained-SiGe layer at low temperatures in a microwave oxygen plasma," *Appl. Phys. Lett.*, vol. 65, pp. 895–897 (see also Vol. 66(12), p.1566, March 20, 1995 issue), 1994.

[41] P. J. Wright, A. Kermani, and K. C. Saraswat, "Nitridation and post nitridation anneals of $SiO_2$ for ultrathin dielectrics," *IEEE Trans. Electron Dev.*, vol. ED-37, pp. 1836–1841, 1990.

[42] M. Bhat, G. W. Yoon, J. Kim, and D. L. Kwong, "Effects of $NH_3$ nitridation on oxides grown in pure $N_2O$ ambient," *Appl. Phys. Lett.*, vol. 64, pp. 2116–2118, 1994.

[43] H. Fukuda, T. Arakawa, and S. Ohno, "Highly reliable thin nitrided $SiO_2$ films formed by rapid thermal processing in an $N_2O$ ambient," *Jap. J. Appl. Phys.*, vol. 29, pp. L2333–L2336, 1990.

[44] W. A. Hill and C. C. Coleman, "A single frequency approximation for interface state density determination," *Solid-State Electron.*, vol. 23, pp. 987–993, 1980.

[45] S. Olsen, A. O'Neill, D. Norris, A. Cullis, K. Fobelets, and H. Kemhadjian, "Impact of virtual substrate quality on performance enhancements in strained-Si/SiGe heterojunction n-channel MOSFETs," *Solid-State Electron.*, vol. 47, pp. 1289–1295, 2003.

[46] S. H. Olsen, A. G. O'Neill, D. J. Norris, A. G. Cullis, S. J. Bull, S. C. Chattopadhyay, K. S. K. Kwa, L. S. Driscoll, A. M. Waite, Y. T. Tang, and A. G. R. Evans, "Thermal oxidation of strained-Si/SiGe: impact of surface morphology and effect on MOS devices," *Mat. Sci. and Eng.*, vol. 109, pp. 78-84, 2004.

[47] S. H. Olsen, A. G. O'Neill, D. J. Norris, A. G. Cullis, N. J. Woods, J. Zhang, K. Fobelets, and H. A. Kemhadjian, "Strained Si/SiGe n-channel MOSFETs: impact of cross-hatching on device performance," *Semicond. Sci. Technol.*, vol. 17, pp. 655–661, 2002.

[48] Q. Xiang, J.-S. Goo, J. Pan, B. Yu, S. Ahmed, J. Zhang, and M.-R. Lin, "Strained silicon NMOS with nickel-silicide metal gate," *IEEE VLSI Tech. Symp. Dig.*, pp. 101–102, 2003.

[49] M. T. Currie, C. W. Leitz, T. A. Langdo, G. Taraschi, E. A. Fitzgerald, and D. A. Antoniadis, "Carrier mobilities and process stability of strained-Si n- and p-MOSFETs on SiGe virtual substrates," *J. Vac. Sci. Technol. B*, vol. 19, pp. 2268–2279, 2001.

[50] C. K. Maiti, L. K. Bera, and S. Chattopadhyay, "Strained-Si heterostructure field effect transistors," *Semicond. Sci. Technol.*, vol. 13, pp. 1225–1246, 1998.

[51] L. K. Bera, S. Mathew, N. Balasubramanian, C. Leitz, G. Braithwaite, F. Singaporewala, J. Yap, J. Carlin, T. Langdo, T. Lochtefeld, M. Currie, R. Hammond, J. Fiorenza, H. Badawi, and M. Bulsara, "Investigation on electrical properties of furnace grown gate oxide on strained-Si," *Thin Solid Films*, vol. 462-463, pp. 85–89, 2004.

[52] L. K. Bera, M. Mukhopadhyay, S. K. Ray, D. K. Nayak, N. Usami, Y. Shiraki, and C. K. Maiti, "Oxidation of strained-Si in a microwave electron cyclotron resonance plasma," *Appl. Phys. Lett.*, vol. 70, pp. 217–219, 1997.

[53] C. G. Ahn, H. S. Kang, Y. K. Kwon, and B. Kang, "Effects of segregated Ge on electrical properties of $SiO_2$/SiGe interface," *Jpn. J. Appl. Phys.*, vol. 37, pp. 1316–1319, 1998.

[54] E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology.* Wiley-Interscience, New York, 1982.

[55] M. Spadafora, G. Privitera, A. Terrasi, S. Scalese, C. Bongiorno, A. Carnera, M. Di Marino, and E. Napolitani, "Oxidation rate enhancement of SiGe epitaxial films oxidized in dry ambient," *Appl. Phys. Lett.*, vol. 83, pp. 3713–3715, 2003.

[56] D. Wang, M. Ninomiya, M. Nakamae, and H. Nakashima, "Electrical characterization of strained-Si/SiGe wafers using transient capacitance measurements," *Appl. Phys. Lett.*, vol. 86, pp. 122111–122113, 2005.

[57] D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt, and G. Timp, "The electronic structure at the atomic scale of ultrathin gate oxides," *Nature*, vol. 399, pp. 758–761, 1999.

[58] G. Wilk, R. Wallace, and J. Anthony, "High-k gate dielectrics: current status and materials properties considerations," *J. Appl. Phys.*, vol. 89, pp. 5243–5275, 2001.

[59] R. M. Wallace and G. Wilk, "Alternative gate dielectrics for microelectronics - special issue," *MRS Bull.*, vol. 327, pp. 186–191, 2002.

[60] K. Rim, E. P. Gusev, C. D'Emic, T. Kanarsky, H. Chen, J. Chu, J. Ott, K. Chan, D. Boyd, V. Mazzeo, B. H. Lee, A. Mocuta, J. Welser, S. L. Cohen, M. Ieong, and H.-S. Wong, "Mobility enhancement in strained-Si NMOSFETs with $HfO_2$ gate dielectrics," *IEEE VLSI Tech. Symp. Dig.*, pp. 12–13, 2002.

[61] C. K. Maiti, S. Chatterjee, G. K. Dalapati, and S. K. Samanta, "Electrical properties of $Ta_2O_5$ gate dielectric on strained-Si," *Electronics Lett.*, vol. 39, pp. 497–499, 2003.

[62] C. K. Maiti, S. K. Samanta, G. K. Dalapati, and S. Chatterjee, "Electrical characterization of $SiO_xN_y/Ta_2O_5$ stacked gate oxides on strained-Si," *IPFA 2003*, pp. 146–150, 2003.

[63] D. G. Schlom and J. H. Haeni, "A thermodynamic approach to selecting alternative gate dielectrics," *MRS Bull.*, vol. 327, pp. 198–204, 2002.

[64] D. J. Hubbard and D. J. Schlom, "Thermodynamic stability of binary oxides in contact with silicon," *J. Mater. Res.*, vol. 11, pp. 2757–2776, 1996.

[65] R. M. Wallace and G. D. Wilk, "Identifying the most promising high-k gate dielectrics," *Semicond. Int.*, 2001.

[66] H. Treichel, A. Mitwalsky, G. Tempel, G. Zorn, D. A. Bohling, K. R. Coyle, B. S. Felker, M. George, W. Ker, A. P. Lane, and N. P. Sandler, "Deposition, annealing and characterization of high-dielectric-constant metal oxide films," *Adv. Mat. for Optics and Electron.*, vol. 5, pp. 163–175, 1995.

[67] M. Johansson, M. Y. A. Yousif, P. Lundgren, S. Bengtsson, J. Sundqvist, A. Harsta, and H. H. Radamson, "$HfO_2$ gate dielectrics on strained-Si and strained-SiGe layers," *Semicond. Sci. Technol.*, vol. 18, pp. 820–826, 2003.

[68] C. K. Maiti, S. Maikap, S. Chatterjee, S. K. Nandi, and S. K. Samanta, "Hafnium oxide gate dielectric for strained-$Si_{1-x}Ge_x$," *Solid-State Electron.*, vol. 47, pp. 1995–2000, 2003.

[69] S. Chatterjee, G. K. Dalapati, S. K. Samanta, and C. K. Maiti, "Electrical properties of $ZrO_2$ films on $Si_{1-x-y}Ge_xC_y$ epitaxial layers," *1st International SiGe technology and Device Meeting (ISTDM 2003), Japan*, pp. 45–46, 2003.

[70] S. Chatterjee, S. K. Samanta, H. D. Banerjee, and C. K. Maiti, "Deposition of high-k $ZrO_2$ films on strained-SiGe layers using microwave plasma," *Electron. Lett.*, vol. 37, pp. 390–392, 2001.

[71] S. Chatterjee, S. K. Samanta, and C. K. Maiti, "Structural characterization and effects of annealing on the electrical properties of stacked $SiO_xN_y/Ta_2O_5$ ultrathin films on strained-$Si_{0.82}Ge_{0.18}$ substrates," *J. Phys. D: Appl. Phys.*, vol. 36, pp. 901–907, 2003.

[72] S. Chatterjee, S. K. Samanta, H. D. Banerjee, and C. K. Maiti, "Electrical properties of stacked gate dielectric ($SiO_2/ZrO_2$) deposited on strained SiGe layers," *Thin Solid Films*, vol. 422, pp. 33–38, 2002.

[73] S. Chatterjee, S. K. Samanta, H. D. Banerjee, and C. K. Maiti, "Metallo-organic compound-based plasma enhanced CVD of $ZrO_2$ films for microelectronic applications," *Bull. Mater. Sci.*, vol. 24, pp. 579–582, 2001.

[74] S. Chatterjee, S. K. Samanta, H. D. Banerjee, and C. K. Maiti, "Effect of the stack layer on the electrical properties of $Ta_2O_5$ gate dielectrics deposited on strained-$Si_{0.82}Ge_{0.18}$ substrates," *Semicond. Sci. Technol.*, vol. 17, pp. 993–998, 2002.

[75] G. K. Dalapati, S. Chatterjee, S. K. Samanta, and C. K. Maiti, "Electrical characterization of low temperature deposited $TiO_2$ films on strained-SiGe layers," *Appl. Surf. Sci.*, vol. 210, pp. 249–254, 2003.

[76] G. K. Dalapati, S. Chatterjee, S. K. Samanta, S. K. Nandi, P. K. Bose, S. Varma, S. Patil, and C. K. Maiti, "Electrical properties of ultra thin $TiO_2$ films on $Si_{1-y}C_y$ heterolayers," *Solid-State Electron.*, vol. 47, pp. 1793–1798, 2003.

[77] C. O. Chui, S. Ramanathan, B. B. Triplett, P. C. McIntyre, and K. C. Saraswat, "Germanium MOS capacitors incorporating ultrathin high-$\kappa$ gate dielectric," *IEEE Electron Dev. Lett.*, vol. 23, pp. 473–475, 2002.

[78] S. K. Ray, C. K. Maiti, S. K. Lahiri, and N. B. Chakrabarti, "Properties of silicon dioxide films deposited at low temperatures by microwave plasma enhanced decomposition of tetraethylorthosilicate," *J. Vac. Sci. Technol. B*, vol. 10, pp. 1139–1150, 1992.

[79] K. Kukli, J. Aarik, A. Aidla, H. Siimon, M. Ritala, and M. Leskela, "In situ study of atomic layer epitaxy growth of tantalum oxide thin films from $Ta(OC_2H_5)_5$ and $H_2O$," *Appl. Surf. Sci.*, vol. 112, pp. 236–242, 1997.

[80] E. Atanassova, T. Dimitrova, and J. Koprinarova, "AES and XPS study of thin RF-sputtered $Ta_2O_5$ layers," *Appl. Surf. Sci.*, vol. 84, pp. 193–202, 1995.

[81] S. Miyazaki, M. Narasaki, M. Ogasawara, and M. Hirose, "Chemical and electronic structure of ultrathin zirconium oxide films on silicon as determined by photoelectron spectroscopy," *Solid State Electron.*, vol. 46, pp. 1679–1685, 2002.

[82] M. Guittet, J. Crocombette, and M. Gautier-Soyer, "Bonding and XPS chemical shifts in $ZrSiO_4$ versus $SiO_2$ and $ZrO_2$: charge transfer and electrostatic effects," *Phys. Rev. B*, vol. 63, pp. 125117–1–125117–7, 2001.

[83] J. Maserjian, G. Peterssion, and C. Svenssion, "Saturation capacitance of thin oxide MOS structures and the effective surface density of states of silicon," *Solid-State Electron.*, vol. 17, pp. 335–339, 1974.

[84] J. Robertson, "Band offsets of wide-band-gap oxides and implications for future electronic devices," *J. Vac. Sci. Technol. B*, vol. 18, pp. 1785–1791, 2000.

[85] C. K. Maiti, S. K. Samanta, G. K. Dalapati, S. K. Nandi, and S. Chatterjee, "Electrical characterization of $TiO_2$ gate oxides on strained-Si," *Microelectronic Eng.*, vol. 72, pp. 253–256, 2004.

[86] C. K. Maiti, G. K. Dalapati, S. Chatterjee, S. K. Samanta, S. Varma, and S. Patil, "Electrical properties of high permittivity $ZrO_2$ gate dielectrics on strained-Si," *Solid-State Electron.*, vol. 48, pp. 2235–2241, 2004.

[87] J.-H. Lee, S. Maikap, D.-Y. Kim, R. Mahapatra, S. K. Ray, Y. S. No, and W. K. Choi, "Characteristics of ultrathin $HfO_2$ gate dielectrics on strained-$Si_{0.74}Ge_{0.26}$ layers," *Appl. Phys. Lett.*, vol. 83, pp. 779–781, 2003.

[88] S. M. Sze, *Physics of Semiconductor Devices.* John Wiley & Sons, New York, 2nd ed., 1981.

[89] W. Harrell and C. Gopalakrishnan, "Implications of advanced modeling on the observation of Poole-Frenkel effect saturation," *Thin Solid Films*, vol. 405, pp. 205–217, 2002.

[90] S. J. Lee, C. H. Choi, A. Kamath, R. Clark, and D. L. Kwong, "Characterization and reliability of dual high-k gate dielectric stack (poly-Si-$HfO_2$-$SiO_2$) prepared by in-situ RTCVD process for system-on-chip applications," *IEEE Electron Dev. Lett.*, vol. 24, pp. 105–107, 2003.

[91] J. W. McPherson, J. Kim, A. Shanware, H. Mogul, and J. Rodriguez, "Trends in the ultimate breakdown strength of high dielectric-constant materials," *IEEE Trans. Electron Dev.*, vol. 50, pp. 1771–1778, 2003.

[92] R. Degraeve, T. Kauerauf, A. Kerber, E. Cartier, B. Govoreanu, Pb. Roussel, L. Pantisano, P. Blomme, B. Kaczer, and G. Groeseueke, "Stress polarity dependence of degradation and breakdown of $SiO_2$/high-k stacks," *Proc. Int. Reliability Phys. Symp.*, pp. 23–28, 2003.

[93] K. Onishi, C. S. Kang, R. Choi, H.-J. Cho, S. Gopalan, R. Nieh, S. Krishnan, and J. C. Lee, "Effects of high-temperature forming gas anneal on $HfO_2$ MOSFET performance," *IEEE VLSI Tech. Symp. Dig.*, pp. 22–23, 2002.

[94] D. Han, J. Kang, C. Lin, and R. Han, "Reliability characteristics of high-k gate dielectrics $HfO_2$ in metal-oxide semiconductor capacitors," *Microelectronic Eng.*, vol. 66, pp. 643–647, 2003.

[95] K.-H. Allers, "Prediction of dielectric reliability from I-V characteristics: Poole-Frenkel conduction mechanism leading to $\sqrt{E}$ model for silicon nitride MIM capacitor," *Microelectron. Reliab.*, vol. 44, pp. 411–423, 2004.

[96] Y.-H. Kim and J. C. Lee, "Reliability characteristics of high-k dielectrics," *Microelectron. Reliab.*, vol. 44, pp. 183–193, 2004.

[97] H. Shang, H. Okorn-Schmidt, K. K. Chan, M. Copel, J. Ott, P. M. Kozlowski, S. Steen, R. M. Sicina, E. Sikorski, S. Cordes, H.-S. P. Wong, E. C. Jones, and W. Haensch, "High mobility p-channel germanium MOSFETs with a thin Ge oxynitride gate dielectric," *IEEE IEDM Tech. Dig.*, pp. 441–444, 2002.

[98] C. O. Chui, H. Kim, D. Chi, B. B. Triplett, P. C. McIntyre, and K. C. Saraswat, "A sub-400$^o$C germanium MOSFET technology with high-$\kappa$ dielectric and metal gate," *IEEE IEDM Tech. Dig.*, pp. 437–440, 2002.

[99] M. L. Lee and E. A. Fitzgerald, "Optimized strained-Si/strained Ge dual-channel heterostructures for high mobility P- and N-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 429–432, 2003.

# 6

# *Heterostructure SiGe/SiGeC MOSFETs*

The increasing demand for wireless communication products is translating into innovation in system architecture, as well as circuits design methodology and techniques. At the core of these approaches is high emphasis in finding the right integrated chip that provides small size, low cost, fewer off-chip components, better integration, and low operating voltage. To serve these needs, III-V semiconductors such as GaAs have been used successfully so far. However, the integration of highly complex digital circuits on GaAs is often prohibitive because of high cost, limited wafer size, processing complexities, and poor yield. The maturity, integration level, and low-cost aspects of Si technology make it attractive for this high-volume marketplace. Besides its superiority at intermediate frequencies, the recent advances in Si/SiGe heterostructure devices and circuits allows it to be competitive with GaAs at microwave and millimeter wave frequencies.

The availability of SiGe-HBT BiCMOS technology [1, 2] with both very high performance HBTs ($f_{max}$ = 300 GHz) and 90-nm gate length CMOS for logic and memory offers the possibility of combining analog and digital components on the same chip, resulting in a new “single chip” architecture. Compressively strained-$Si_{1-x}Ge_x$ provides an avenue to improve hole mobility and thus increase p-MOSFETs drive current. The first report of enhanced mobility $Si_{1-x}Ge_x$ p-MOSFET was by Nayak et al. [3]. The SiGe heterostructure field-effect transistors (SiGe HFETs) is presently catching up in high-frequency performance due to higher mobilities in 2D electron and 2D hole within strained-Si channel [4, 5]. A 90% mobility enhancement for $Si_{0.7}Ge_{0.3}$ on SIMOX over identical Si control devices was reported.

The enhancement electron mobility in strained-Si layers compared to bulk-Si has been a subject of great interest, and offers the promise of high-speed devices based on Si technology. 2D electron gases (2DEGs) were created at a Si/SiGe heterointerface in thin tensile strained-Si layers deposited on layers of relaxed-SiGe alloy, using Si epitaxy. Modulation doping was used to reduce carrier scattering to form a structure similar to 2DEGs in III-V material systems, separating the doped region from the conduction channel region. Si/SiGe 2DEGs have been found to have high electron mobilities at a low temperature and have opened up the possibility of improving the room temperature performance of Si devices.

The $Si_{1-x}Ge_x$ device is in general a buried-channel device because a high-quality gate oxide cannot be grown on $Si_{1-x}Ge_x$; thus a Si cap is necessary on

top of the $Si_{1-x}Ge_x$ to enable the growth of a thermal gate oxide. In order to maximize performance in buried-channel $Si_{1-x}Ge_x$ p-MOSFETs, the Si cap needs to be as thin as possible, though there are trade-offs. Thicker Si caps degrade gate capacitance, but very thin caps lead to increased surface roughness scattering [6]. SOI MOSFETs show a performance advantage over conventional Si MOSFETs because of improved device isolation, reduced parasitic capacitance, lower power consumption, better subthreshold swing, and higher current drive. In order to get higher hole confinement in the strained-SiGe-channel, SiGe p-MOSFETs have been demonstrated on a silicon-on-insulator structure using SIMOX [7]. Improved quantum confinement at the SiGe-well on SIMOX substrate compared to SiGe on Si substrate was reported using photoluminescence (PL) measurement.

The growth of high-quality $Si/Si_{1-x}Ge_x/Si$ structure on SIMOX has been demonstrated. The SOI structure consists of a Si-cap layer on a strained-$Si_{1-x}Ge_x$ on buried-oxide-layer. From the viewpoint of device technology, however, the biggest concern is the floating-body effect, which is known to manifest itself as lowering of abnormal subthreshold slope. To suppress the floating body effect in SOI MOSFETs, Yoshimi et al. [8] have employed bandgap engineering using a SOI SiGe structure. Mathew et al. [9] demonstrated the enhanced low-field mobility, transconductance, and noise figure performance for SiGe MOSFETs using silicon-on-sapphire technology. Reported room- and low-temperature hole mobility enhancement using SiGe-channel was seen to vary from 30% to 70% (depending upon Ge concentration) over that of Si. Yeo et al. [10] reported a SOI SiGe p-channel MOSFET with 50-nm channel length. Enhanced drive current (70%) was reported [10] due to the introduction of $Si_{0.7}Ge_{0.3}$ on SOI structure.

In chapter 4 on electronic properties of strained-Si, it has been shown that electron or hole confinement in structures (for n-HFET or p-HFET) require more complex growth techniques for the strained-Si on relaxed thick SiGe layers and are limited in terms of processing thermal budget. In contrast, the p-HFET is more easily realized, since it involves the growth of strained-$Si_{1-x}Ge_x$ epitaxial films on a Si substrate. In this device, the $Si_{1-x}Ge_x$ quantum well acts as a channel for holes between the source and drain regions of the device as shown in Fig. 6.1. Improved electrical characteristics of this device over the conventional surface channel Si p-MOSFET are results of improved carrier transport, quantum confinement, and buried-channel operation. p-HFETs provide quantum confined carrier conduction with high carrier mobility, which is critical for high-frequency Si-based integrated circuits.

A number of variants of the basic SiGe-HFET have been reported, such as p-HFETs built on SOI substrates and SiC/SiGeC channel devices. Vertical SiGe and SiGeC p-HFETs are also attractive for ultrashort channel devices because the channel length is determined by the thickness of an epitaxial layer and not by the lithography resolution. Poly-$Si_{1-x}Ge_x$ has shown a great potential as a gate material due to its tunable work function, process compatibility, and favorable electrical properties, such as low sheet resistance, and high dopant

**TABLE 6.1**
Several important material parameters for C, Si, and Ge.

| Element | Atomic No. | Lattice Const. (Å) | Lattice mismatch to Si (%) | Structure, Bond length (Å) | Bandgap (eV) |
|---|---|---|---|---|---|
| C diamond | 6 | 3.567 | -34.32 | diamond, 1.55 | 5.48 |
| Si | 14 | 5.431 | 0 | diamond, 2.34 | 1.11 |
| Ge | 32 | 5.657 | + 4.17 | diamond, 2.44 | 0.664 |

activation rate.

In this chapter, a review on the present status of silicon heterostructure field-effect transistors in the SiGe and SiGeC material systems is presented. The physics and modeling of submicron p-HFETs are explored, using numerical simulation to determine the potential applications in ULSI circuits. The key design issues such as Ge mole fraction, gate oxide thickness and choice of gate contact material have been considered in detail. The choice of the cap layer thickness for a buried SiGe channel is an important issue, having bearing on the performance of a p-HFET, and due consideration is given.

Design and simulation issues for advanced SiGe p-MOSFETs have been investigated. Charge distribution in the Si-cap and SiGe-well of $Si_{1-x}Ge_x$ buried-channel MOS capacitor structures has been calculated using analytical model and 1D Poisson solver. Semi-analytical model, 1D Poisson solver and Silvaco-ATLAS tools have been used to optimize the $Si_{1-x}Ge_x$ channel p-MOSFETs device performance for the modulation-doped structure. Results of simulation of strained-Si HEMTs investigated by 2D device simulator using Silvaco-ATLAS are also presented.

## 6.1 SiGe/SiGeC: Material Parameters

Several material parameters for strained-Si and relaxed-SiGe are needed (see Table 6.1) for the device design and modeling, viz., the band offset encountered in strained layers, doping-induced bandgap narrowing, and the carrier mobility. Although Ge is completely miscible in Si, epitaxial films of $Si/Si_{1-x}Ge_x$ can be either in biaxial compression or tension, depending on the lattice constant of the substrate on which it is grown because of their larger lattice mismatch (4.2%). Silicon strained in biaxial tension results in a type-II band offset [11], and fairly large band offset (in order of 100 meV) in both the valence and conduction bands, relative to the relaxed-$Si_{1-x}Ge_x$ substrate. This allows tailoring of the band structure to confine both holes and electrons. On the other hand, under biaxial compression strained-$Si_{1-x}Ge_x$ produces a type-I band offset, i.e., the entire band offset occurs in the valance band while

band offset in conduction band is very small. Under biaxial strain, both the strained-Si and strained-$Si_{1-x}Ge_x$ bandgap are reduced compared to its bulk value, which has empirically been determined as

$$E_g^{Si}(x) = 1.11 - 0.4x, \quad E_g^{SiGe}(x) = 1.11 - 0.74x \tag{6.1}$$

for strained-Si and strained-SiGe, respectively, where x is the the Ge concentration in $Si_{1-x}Ge_x$ substrate. The conduction band discontinuity, $\Delta E_c$, for strained-Si grown on a relaxed-$Si_{0.7}Ge_{0.3}$ buffer, has been reported to be 160 meV [12]. This amount of conduction band discontinuity is sufficient to confine electrons in the strained-Si channel at 300 K and has been utilized for the fabrication of strained-Si channel HEMT.

Although several doping-dependent bandgap narrowing models have been proposed for silicon [13, 14], very little information is available in the literature for strained-Si. In simulation, we assume that the bandgap narrowing due to heavy doping in strained-Si layers to be the same as in silicon. This assumption is reasonable for doping concentrations up to approximately 1x$10^{17}$ $cm^{-3}$. Also, the doping-induced bandgap narrowing model [15] for Si is used for SiGe layers. This approach has the advantage that any difference in the simulation of Si and strained-Si HEMT can be unambiguously attributed to heterojunction action, rather than differences in model parameters.

The other important model parameter in a device simulation is the carrier mobility. It is necessary to take into account the enhancement mobility in strained-Si. In strained-Si, the strain releases the six-fold degeneracy of the conduction band and is split into two lowered and four raised $\Delta$-valleys, while reverse occurs in strained $Si_{1-x}Ge_x$. The lower $\Delta_\perp$-valley exhibits a lower transport mass and, due to the splitting of the valley, intervalley scattering is reduced. As a result, the electron mobility is increased. In simulation, the low-field electron mobility for strained-Si channel has been modeled following Vogelsang and Hofmann [16], who calculated the electron mobility for different values of x. For doping dependence of strained-Si mobility, the Arora mobility model for Si is modified by using an analytical expression involving Ge fraction, x, of the substrate as follows:

$$\mu_p = \mu_{min} T_n^{\alpha} + \frac{\mu_o T_n^{\beta}}{1 + (\frac{N}{N_{ref}})^{\delta} T_n^{\gamma}} \tag{6.2}$$

where $T_n = T/300$, and $\mu_{min}$ and $\mu_o$ can be fitted as

$$\mu_{min} = 44 - 20x + 850x^2; \qquad \mu_o = 400 + 29x + 4737x^2 \tag{6.3}$$

The temperature dependence parameter $\alpha$, $\beta$, and $\gamma$ are assumed same as Si and are shown in Table 6.2. On the other hand, the incorporation of Ge introduces alloy scattering in strained-SiGe layers and the strain affects the mobility through the reduction in the effective masses. The alloy-scattering-limited mobility components along perpendicular and parallel to the growth direction for coherently strained-$Si_{1-x}Ge_x$ is given by, respectively, [17]

**TABLE 6.2**
Hole mobility parameters for strained-SiGe.

| $N_{ref}$(cm$^{-3}$) | $\alpha$ | $\beta$ | $\gamma$ | $\delta$ |
|---|---|---|---|---|
| 2.35×10$^{17}$ | -0.57 | -2.23 | -2.546 | 0.90 |

$$\mu_{\perp}^{alloy} = \frac{5.5 \times 10^{18} \cdot T}{22.0 \cdot N_c \cdot x(1-x) \cdot m_t^2} \tag{6.4}$$

$$\mu_{\parallel}^{alloy} = \frac{5.5 \times 10^{18} \cdot T}{4.0 \cdot N_c \cdot x(1-x) \cdot m_l^2} \tag{6.5}$$

where $N_c$ is the effective density of states for Si and x is the Ge mole fraction. It is noted that the alloy mobility decreases with increasing Ge content. The parallel component $\mu_{\parallel}^{SiGe}$ and perpendicular component $\mu_{\parallel}^{SiGe}$ of electron mobility in SiGe can thus be obtained by simply using Mathiessen's rule using the parallel and perpendicular to the growth plane Si mobility. The individual component of electron mobility of Si is expressed as [17]

$$\mu_{\perp}^{si} = \frac{3\mu_{si}}{\frac{m_t}{m_l} + 2}, \quad \mu_{\parallel}^{si} = \frac{3\mu_{si}}{2(\frac{m_l}{m_t}) + 1} \tag{6.6}$$

where $m_l$ and $m_t$ are longitudinal and transverse density of state masses in Si and $\mu_{Si}$ is the electron mobility in Si and is given by Eqn. 6.6, where $\mu_{min}$ = 44 and $\mu_o$ = 400. Finally, the total electron mobility for the growth plane of strained SiGe for Ge fraction greater than 20% can be approximated as

$$\mu_{xx}^{SiGe} = \frac{\mu_{\perp}^{SiGe} + \mu_{\parallel}^{SiGe}}{2} \tag{6.7}$$

The hydrodynamic model used in simulation includes the effects of carrier heating by the self-consistent solution of the drift-diffusion and energy balance equations. The low-field mobility is in turn determined by solving the homogeneous carrier temperature equation:

$$\mu E_{eff}^2 = \frac{3K}{2q} \cdot \frac{T_n - T_L}{\tau_e} \tag{6.8}$$

where $\mu$ is carrier mobility, $T_n$ is the carrier temperature, $T_L$ is the lattice equilibrium temperature, and $\tau_e$ is the energy relaxation time. The effective electric field can then be used in the Caughey-Thomas expression for mobility [18]. However, the saturation velocity of strained-Si has been reported to be approximately the same as that of bulk-Si [19, 20].

The dielectric constant of a material determines how strongly it adjusts its electron distribution to an applied electric field. In case of $Si_{1-x}Ge_x$ alloys, the effect depends on the polarizability of the atoms of the materials, and

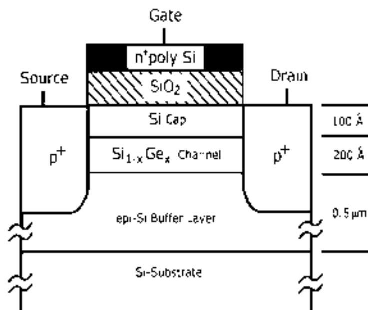



**FIGURE 6.1**
Device structure for a typical SiGe-channel p-HFET fabricated in a strained-SiGe layer with a Si-cap layer. Layer thicknesses shown are typical.

the crystal structure formed by the $Si_{1-x}Ge_x$ alloy is essentially the same as that of a pure Si or pure Ge crystal. However, the dielectric constant of strained-Si is assumed to be 11.9 and for $Si_{1-x}Ge_x$, it is obtained from a linear interpolation of dielectric constant for that of Ge and Si, and is given by expression

$$\epsilon_{SiGe}(x) = \epsilon_S i + x(\epsilon_{Ge} - \epsilon_{Si}). \tag{6.9}$$

## 6.2 SiGe Hetero-FETs: Structures and Operation

Currently, high-performance bipolar and BiCMOS SiGe technology are being extensively used for RFIC applications. The major advantage of SiGe is its compatibility with conventional CMOS technology and both field-effect and bipolar devices may be fabricated on a single Si chip [21]. For MOSFETs, various types of channel have been proposed; surface channel in which carriers propagate along the surface, while in the buried-channel device, added advantages of improved immunity to hot-carrier degradation over the surface-channel device are obtained. The buried-channel operation has been known to yield a higher carrier mobility due to reduction in the field and avoids surface scattering. However, buried- and surface-channel MOSFETs behave similarly in long channel regime. A substantial difference (threshold voltage lowering and poor turnoff characteristics) is observed in device operation in short-channel region.

The SiGe p-HFET with a general structure similar to a conventional MOSFET is shown in Fig. 6.1. It has an $n^+$- (or $p^+$)-polySi or poly-SiGe gate over a thin gate oxide, with $p^+$ source/drain regions in a n-type body. The

main distinctive features are the buried SiGe layer and the optional $p^+$ doping spike ($\delta$-doping) located below it. The SiGe layer actually constitutes a subsurface quantum well channel for holes between the source/drain regions. It is required to be buried below a Si-cap because a high-quality gate oxide directly on SiGe using thermal techniques is difficult to obtain. If such a direct oxidation of SiGe is attempted, the Si is preferentially oxidized, leading to a pile-up of Ge at the SiGe/$SiO_2$ interface [22, 23]. Plasma-enhanced chemical vapor-deposited oxides do not have significantly lower interface state densities, since the initial stages of this process consist of oxide growth and not deposition [24]. The interface state density of such oxides is greater than $10^{12}$ $cm^{-2}eV^{-1}$, a figure unsuitable for proper operation of a field-effect device. Stoichiometric oxides formed directly on SiGe using low temperature plasma techniques have been reported [25, 26] and microwave/ECR plasma-grown oxides have been used for device fabrication [27].

The various studies have shown that the formation of low-mobility surface parasitic channel in conventional Si/$Si_{1-x}Ge_x$ heterojunction p-MOSFETs greatly reduces their performance advantage over conventional p-MOSFETs [3]. To reduce the population of the parasitic channel, Si/$Si_{1-x}Ge_x$ heterojunction p-MOSFETs with a buried $\delta$-doping layer have been proposed by some researchers. Sadek et al. [28] have demonstrated that the $\delta$-doping layer helps to reduce the parasitic channel to a minimum and the resulting design has a big advantage over conventional MOSFETs in drain current driving ability. In particular, the effects of layer thickness, $\delta$-doping dose and potential well depth on device performance are very important.

Several research groups have fabricated SiGe-channel p-HFETs, mostly using conventional Si processing technology, and performance enhancement compared to bulk-Si devices has been reported [3, 6, 7, 27, 29, 30, 31]. In some designs the Ge profile was graded to optimize the hole confinement and modulation doping ($\delta$-doping) was used to adjust the threshold voltage. Such a SiGe-channel p-HFET is also known as a modulation doped SiGe p-MOSFET (MODMOS) [6].

Other possible configurations of MOSFETs using strained-$Si_{1-x}Ge_x$ include a single layer of thin strained-Si grown on top of the relaxed-$Si_{1-x}Ge_x$ buffer layer. This layer is oxidized to form gate oxide. The structure shows the type-II band offsets and has several advantages over the more common type-I band alignment, as a large band offset is obtained in both the conduction and valence bands, relative to the relaxed $Si_{1-x}Ge_x$ layer [32]. This is favorable for both the electron and hole confinement, making it useful for both n- and p-type devices for strained-Si/SiGe-based CMOS technology [33, 34]. To reduce the carrier population in the surface channel, Si/$Si_{1-x}Ge_x$ heterojunction p-MOSFET with a buried $\delta$-doping layer, separated by a Si spacer layer, have been proposed by several researchers [28, 35, 36]. The doping spike performs two functions viz., (a) it creates a retarding electric field for holes at zero gate bias to suppress source/drain leakage current (threshold adjust), and (b) it provides holes for the $Si_{1-x}Ge_x$ channel for improved device transconductance.

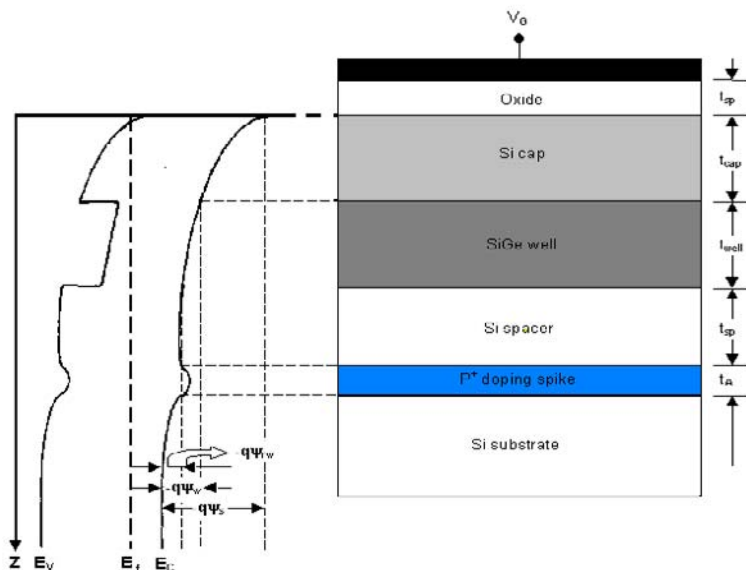



**FIGURE 6.2**
A cross-sectional schematic of a $Si_{1-x}Ge_x$ quantum well MOS structure and band diagram for an arbitrary negative gate bias. After S. Maikap, *Ultra-thin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

In the following, we present a comprehensive design trade-off study as functions of device parameters for $p^+$ $\delta$-doped $Si/Si_{1-x}Ge_x$ heterojunction MOSFETs as shown in Fig. 6.2. We shall define design constraints and performance trade-offs, such as short-channel immunity, off-state leakage and on-state driving ability, in terms of device parameters. Device performance then is examined as a function of layer thicknesses, potential well depth, and $\delta$-doping spike. Finally, a design space construction example is used as a summary of the design procedure. Silvaco-TCAD tool is used to investigate the design and fabrication of submicron p-$Si_{1-x}Ge_x$ heterostructure MOSFETs with and without $p^+$-$\delta$ doping spike.

An ideal MOS structure is shown in Fig. 6.2, which consists of the following layer sequence: Si substrate; $p^+$-doping spike; Si spacer layer; SiGe quantum well; Si-cap layer; oxide and gate metal. It is assumed that the semiconductor region is doped uniformly throughout with $N_d$ donors/cm$^3$, with the exception of the $p^+$-doping spike which is doped with $N_A$ acceptors/cm$^3$, and thin enough so that it is fully depleted at zero gate bias. The band diagram shown in Fig. 6.2 is for an arbitrary negative gate bias condition and defines the notation used throughout this section. The treatment presented below utilizes quantum mechanical description for mobile charge distributions in the Si-cap and in the SiGe layers, denoted by $Q_{cap}$ and $Q_{well}$, respectively. The gate

potential is related to the internal charges and potential viz.:

$$V_g = \psi_s + V_{ox} + V_{FB} \tag{6.10}$$

where $\psi_s$ is the internal electrostatic potential at the Si/SiO$_2$ interface, $V_{ox}$ is the voltage drop across the oxide due to free carrier and depletion charges within the device, and $V_{FB}$ is the flatband voltage, and the total charge is given by

$$Q_{tot} = Q_{cap} + Q_{well} + Q_{dep} \tag{6.11}$$

where $Q_{cap}$ = -q$N_d t_{cap}$ and $Q_{well}$ = -q$N_d t_{well}$.

In structures that do not contain any doping spike and a spacer layer, the carrier distribution computation begins approximately at the flatband voltage. For a device that does contain $p^+$-$\delta$ doping and a spacer layer, the layer is fully depleted and the doping spike can be treated as a sheet of negative charge, similar to the fixed oxide charge in a conventional Si MOS structure.

To find the free carrier density in the SiGe layer ($Q_{well}$), the region is treated as a quantum well for holes due to the observation that most of the bandgap discontinuity between strained-SiGe and Si resides in the valence band. Using the infinite well approximation, the eigen energy:

$$E_n = \frac{\pi^2 h^2 n^2}{2m^*_{\|} t^2_{well}}, (n = 1, 2, 3, .....) \tag{6.12}$$

with the carrier wave function, $\phi_n^{(0)}$, given by:

$$\phi_n^{(0)} = \sqrt{\frac{2}{t_{well}}} sin\left[\frac{n\pi x}{t_{well}}\right] \tag{6.13}$$

where the effective mass, m$^*$, refers to the longitudinal effective mass since quantization is occurring along the z direction [37]. The areal hole density ($p_{well}$) is then obtained by multiplying the two-dimensional density of states with the Fermi integral of order zero:

$$p_{well} = \frac{2m^*_d}{\pi\hbar^2} \int_{-\infty}^{E_l} \frac{dE}{1 + exp[(E - E_f)/kT]} \tag{6.14}$$

where $m_d{}^*$ refers to the density-of-states effective mass of the heavy hole band at the Brillouin zone center [38]. A hole degeneracy factor of two is used in the above rather than the conventional value of four since the strain in the SiGe layer lifts the degeneracy at the zone center between the heavy hole and light hole bands. The foregoing derivation for $p_{well}$ is only valid for the condition where there is no potential variation across the well (i.e., $\psi_w = \psi_{rw}$). With an applied gate bias, the eigen energies change (Stark shift) and can be approximated to first order using nondegenerate perturbation theory as:

$$E_n = E_n^{(0)} + H_{nn} \tag{6.15}$$

where $E_n{}^{(0)}$ are unperturbed eigen energies and $H_{nn} = (|\psi_w\text{-}\psi_{rw}|)/2$, represents the $n^{th}$ level perturbation matrix element. The areal electron density in the SiGe layer can be computed for a given hole density from the law of mass action:

$$n_{well} = \frac{(n_i{}^{SiGe})^2 t_{well}{}^2}{p_{well}} \tag{6.16}$$

where $n_i{}^{SiGe}$ is the intrinsic carrier concentration in $Si_{1-x}Ge_x$. The net areal mobile charge density in the SiGe layer is the difference between the areal hole and electron charge densities, $Q_{well} = q(p_{well}\text{-}n_{well})$.

For larger band bending at the $Si/SiO_2$ interface, the Boltzmann approximation fails to account for the degenerate buildup of carriers. Under these conditions, the Si-cap be treated as a quasitriangular potential well for holes. In this case, the $Si/SiO_2$ interface is treated as an infinite potential barrier, and eigen energies within the triangular well are approximately given by [39]:

$$E_n = \left[\frac{\hbar^2}{2m^*}\right]^{1/3} \left[\frac{3\pi\alpha}{2}(n + \frac{3}{4})\right]^{2/3} \tag{6.17}$$

where $\alpha = q(|\psi_s\text{-}\psi_w|)/t_{cap}$ is related to the slope of the triangular well. Similar to the two-dimensional formulation used, $p_{cap}$ is found to be:

$$p_{cap} = \frac{4kTm_d^*}{\pi\hbar^2} \sum_{n=0}^{i} ln\{1 + exp\left[\frac{(E_n - E_f)}{kT}\right]\} \tag{6.18}$$

where $m_d{}^*$ refers to the density-of-states hole effective mass for Si and the factor of four accounts for the degeneracy between the heavy and light hole bands at the zone center. The sufficient confinement is defined at the calculation index where the difference, $|\ q\psi_s\text{-}q\psi_w\ |$, exceeds the ground state energy ($n = 0$). Carrier distribution in the Si-cap and SiGe-well as a function of gate bias depends on the charge distribution in the heterolayers using electrostatic potential calculation.

In the following, the charge distribution in the surface (Si-cap) and buried SiGe-channel of $Si_{1-x}Ge_x$ p-MOSFETs under various gate biases using semi-analytical model is presented. The model is based on analytical, quantum mechanical descriptions for hole charge in the SiGe quantum well and Si-cap, and requires no ad-hoc fitting parameters. The charge screening effect that limits the maximum number of charge carriers in the SiGe layer is explicitly revealed using this formulation.

The critical design parameters for SiGe-channel p-MOSFETs are the choice of gate material, the method of threshold voltage adjustment, the SiGe profile in the channel, the Si-cap and well thickness, and gate oxide thickness. The key design parameter of a MOSFET is the device transconductance. This

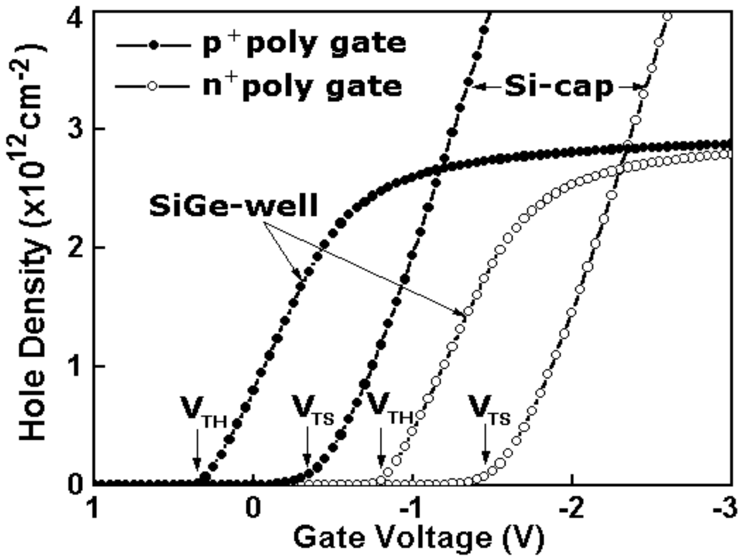



**FIGURE 6.3**
Hole distributions using 1D Poisson's solver as a function of gate voltage for n$^+$- and p$^+$-poly-gate devices. After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

can be enhanced by maximizing the number of high-mobility holes confined to the SiGe channel while minimizing the density of low-mobility holes, which flow at the Si/SiO$_2$ interface. $\delta$-doping spike has been used to enhance the transconductance and better control of threshold voltage adjustment of the device.

The choice of gate material has a significant effect on the turn-on characteristics of the p-MOSFETs. [6]. In Fig. 6.3, the hole distribution in SiGe channel and Si cap layer as a function of gate voltage is given for n$^+$- and p$^+$-gate Si$_{1-x}$Ge$_x$ quantum well p-MOSFETs using 1D Poisson's solver. The baseline structure parameters are used to calculate the hole distribution in Table 6.3.

The hole confinement shows identical nature for both p$^+$- and n$^+$-poly-gate Si$_{1-x}$Ge$_x$ p-MOSFETs. The threshold voltages are different due to different work functions. The amount of shift is equal to Si bandgap. But, the threshold voltage for SiGe MOSFET is lower than that of Si MOSFETs. It may be explained as due to the introduction of energy band discontinuity in the valence band of the Si/Si$_{1-x}$Ge$_x$ interface. Analytically, the SiGe threshold voltage can be expressed as [6]

**TABLE 6.3**
Parameters for baseline structure. After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

| Parameter | Value |
|---|---|
| Work function, $\phi_g$ (eV) | $n^+$ poly |
| Oxide thickness, $t_{ox}$ (nm) | 8 |
| Cap layer thickness, $t_{cap}$ (nm) | 5 |
| Well thickness, $t_{well}$ (nm) | 10 |
| Ge mole fraction, x | 0.3 |
| Channel doping, $N_d$ ($cm^{-3}$) | $1\times10^{16}$ |

$$V_t(SiGe) = V_t(Si) - Q_{channel}/C_{ox} + Q^t_{channel}/C_{eff} - \Delta E_v/q \qquad (6.19)$$

where $V_t$(Si) is defined as the onset of strong inversion in a Si MOSFET. $Q_{channel}$ is the charge in the depletion region of a Si MOSFET. $C_{ox}$ is the oxide capacitance. $Q^t{}_{channel}$ is the background charge in the SiGe channel and in the depleted region underneath the SiGe channel. $C_{ox}$ is replaced by $C_{eff}$, which is the series combination of the oxide capacitance and the Si cap capacitance. $\Delta E_v$ is the valence band offset at the Si/SiGe interface. The bandgap narrowing is approximately 74 meV per 10% Ge introduced and 97% of this offset occurs in the valence band. This difference in threshold voltage between Si and SiGe MOSFET is valid provided that the hole density in the cap layer is negligible at turn-on voltage. This implies the use of thin Si cap layer.

The holes are mostly confined to the quantum well at the buried-channel threshold voltage ($V_t$). If a $p^+$ polysilicon gate is used with a buried-channel p-MOSFET, then the device operates in depletion mode, that is, it is normally on. However, if an $n^+$-polySi gate is used, then the device threshold voltage is shifted negatively by approximately 1 V (silicon's bandgap). Consequently, the buried-channel device operates in enhancement mode with the buried-channel carrier transport properties determining the MOSFET electrical characteristics at low-gate biases. When the gate voltage is decreased to about -1.5 V for $n^+$-gate and -0.5V for $p^+$-gate polysilicon, the number of holes in the parasitic Si cap layer exceeds those in SiGe quantum well. As the gate voltage is decreased further, the hole concentration in the Si cap layer increases rapidly, while the hole concentration in the SiGe channel remains almost constant. This is because the holes at the $Si/SiO_2$ interface shield the vertical field, and any additional decrease in gate voltage appears only across the gate oxide and the charge screening effect.

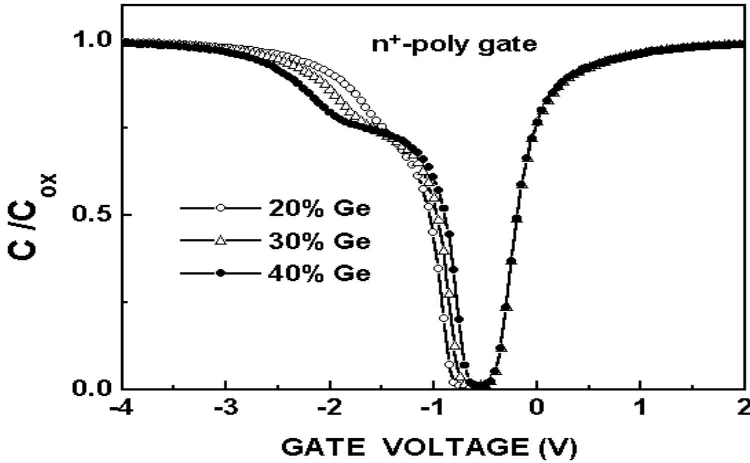



**FIGURE 6.4**
Low-frequency (LF) C-V characteristics of a SiGe-MOS capacitor as a function of Ge composition. After S. K. Samanta et al., *Proc. IWPSD-2003*, 2003(540-542).

For an optimal epitaxial structure, the crossover voltage should occur at a large negative gate bias and with a high hole density in the SiGe channel. To promote high-speed operation, the p-MOSFET should be operated under bias conditions where the hole population in the SiGe well exceeds that in the Si-cap. The crossover voltage for $n^+$-gate is higher than that of $p^+$-gate MOSFET. Hence, a $n^+$ gate design is favorable for the p-MOSFET devices because it promotes SiGe quantum well operation for low gate biases.

Ge mole fraction in the SiGe layer plays a critical role in determining the crossover voltage. A large valence band offset at the top Si/SiGe interface results in a better hole confinement [3]. Strain relaxation issues set the upper limit on the amount of band offset that can be achieved by the incorporation of Ge. For a fixed amount of strain, it is advantageous to minimize the thickness of SiGe layer and increase the Ge mole fraction.

Simulated low-frequency C-V characteristics of $Si_{1-x}Ge_x$ (x = 0.2, 0.3, and 0.4) MOS capacitors using 1D Poisson solver is shown in Fig. 6.4. The plateau in accumulation shows the hole confinement at the $Si/Si_{1-x}Ge_x$ interface. The hole confinement increases with Ge content as the valence band offset increases [40]. As the structure is biased more negatively, an accumulation layer in the Si-cap forms and the capacitance of the structure approaches the oxide capacitance. The accumulation layer carrier population in the surface channel (Si-cap) and buried channel (SiGe-well) have been extracted by integrating the carrier profiles across the total depths of these layers. The enhanced hole population in $Si_{1-x}Ge_x$ with increased Ge concentration (x) is illustrated in

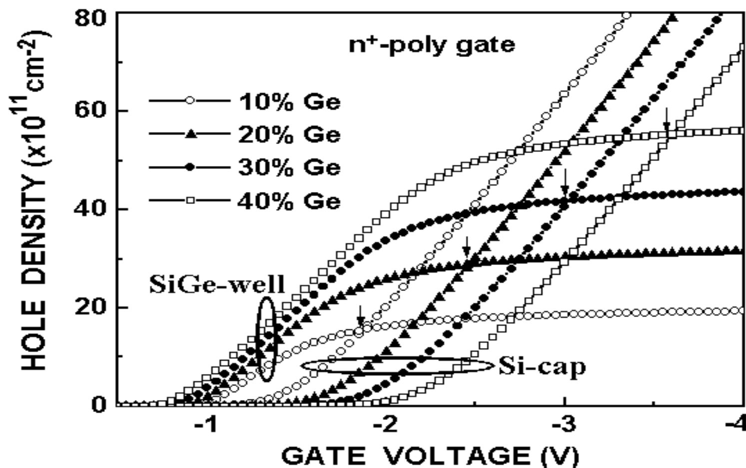



**FIGURE 6.5**
Hole concentration in Si-cap and SiGe-channel as a function of Ge mole fraction in a SiGe p-MOSFET. After S. K. Samanta et al., *Proc. IWPSD-2003*, 2003(540-542).

Fig. 6.5 where the $n^+$-polySi gate MOSFET crossover voltage is shown to increase (arrows) with x. The crossover voltage is linearly proportional to the Ge concentration. As expected, more carriers flow in the Si-cap layer as the gate voltage increases and the device becomes a surface-channel one. In order to ensure that the majority of holes flow in the SiGe-channel for gate voltages equal to or higher than -3.0 V, the Ge percentage must be above 30% [41].

As in dual-channel MOSFETs, control of the strained-Si cap thickness is critical to the operation of MOSFETs. The Si cap thickness on the p-MOSFET can be extracted from a combination of C-V measurements and simulations, as illustrated in Fig. 6.6. As is consistent with type-II band line-ups, for the thicker Si cap devices, the inversion layer forms in the buried SiGe at low gate overdrives and in the strained-Si cap layer at higher overdrives. For the 1-nm thick cap device, the Si cap is too thin to accommodate the holes, and inversion occurs primarily in the SiGe layer, resulting in a slight decrease in the maximum capacitance.

Simulated DC output characteristics of $Si_{1-x}Ge_x$ (x = 0.3; W/L = 10$\mu$m/ 1$\mu$m) devices are shown in Fig. 6.7 for $V_g$ = -2.0, -3.0, and -4.0 V. The drain current increases (compared to control-Si) as the Ge content is increased because of the enhancement of hole mobility [42]. The drain saturation current is found to be -1.56 mA and -3.23 mA for Si and SiGe (x = 0.3), respectively, at $V_d$ = -2.0 V and $V_g$ = -4.0 V. The drain saturation current is 107% higher than that of the Si devices in agreement with the results reported by Sadovnikov et al. [43].

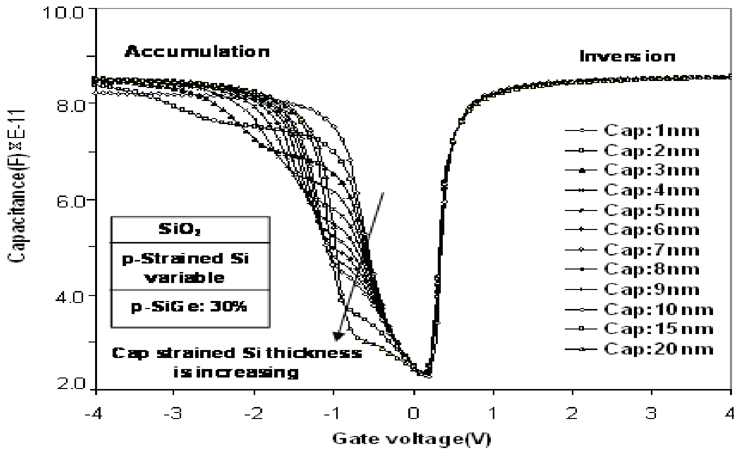



**FIGURE 6.6**
Effect of strained-Si cap layer thickness on carrier confinement at the heterointerface: doping is considered to be p-type. Source: S. Chattopadhyay, unpublished data.

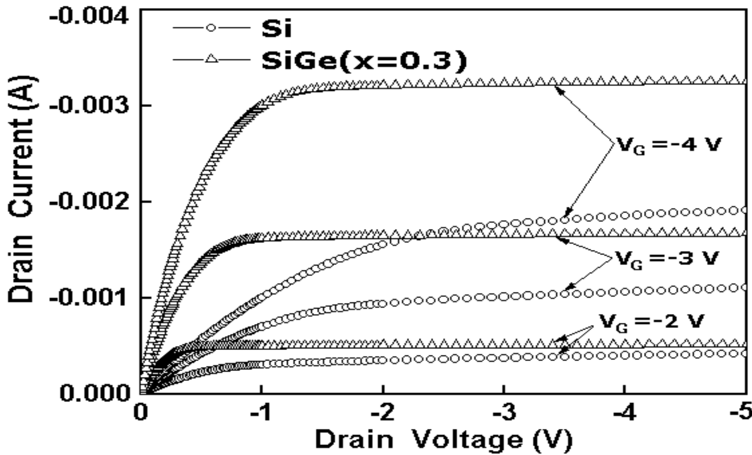



**FIGURE 6.7**
Simulated output characteristics of $Si_{1-x}Ge_x$ and Si channel devices (W/L = 10$\mu$m/1$\mu$m). After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

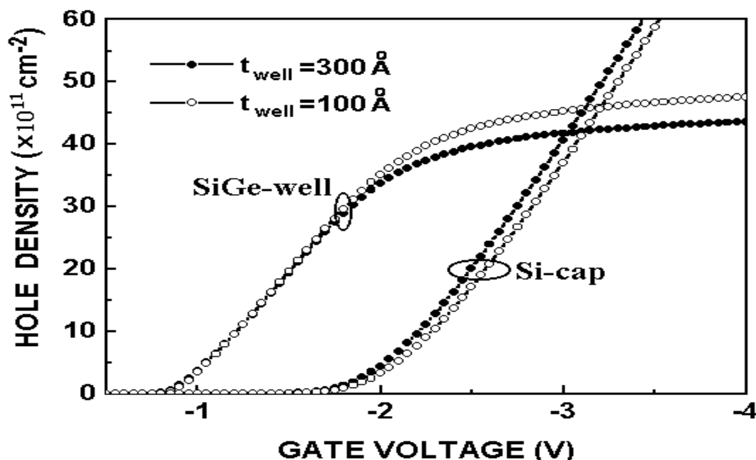



**FIGURE 6.8**
Simulated hole distribution in the SiGe-channel as a function of gate voltage with different SiGe-well thicknesses. After S. K. Samanta et al., *Proc. IWPSD-2003*, 2003(540-542).

1D Poisson simulation results for the sheet concentration of holes in the Si-cap and in the $Si_{1-x}Ge_x$ well as a function of gate voltage for two different well thicknesses (100 Å and 300 Å) are shown in Fig. 6.8. The crossover voltage increases nominally with a decrease in well thickness. The peak transconductance is weakly dependent on SiGe quantum well thickness as shown in Fig. 6.9. The simulation results show that the drive current is optimum when the SiGe layer thickness is in the range 100-300 Å. When SiGe layer is thinner than the normal inversion layer thickness, the Si layer beneath the SiGe layer cannot be inverted due to the Si/SiGe valence band offset. The SiGe well loses some carriers, compared to the case of a thicker SiGe layer, and therefore, the drive current decreases. On the other hand, when the SiGe layer is too thick, the average distance of inversion carriers from the gate becomes larger, and again a smaller drive current results.

Fig. 6.10 presents the 1D Poisson simulation results for hole density in the $Si_{1-x}Ge_x$ layer and in the Si-cap as a function of gate voltage for different cap layer thicknesses. The hole density increases in the cap significantly with an increase in cap layer thickness. The reduction of cap thickness greatly increases the carrier concentration in the $Si_{1-x}Ge_x$ channel. So, one way to increase the concentration of holes in the $Si_{1-x}Ge_x$ layer over that in the Si-cap is to reduce the cap thickness as much as possible. It may be noted that a cap thickness of only 50 Å is enough to support an inversion layer, hence the charge screening problem will still be present in such a structure [44]. Consequently, the device should be operated at low gate voltages where the

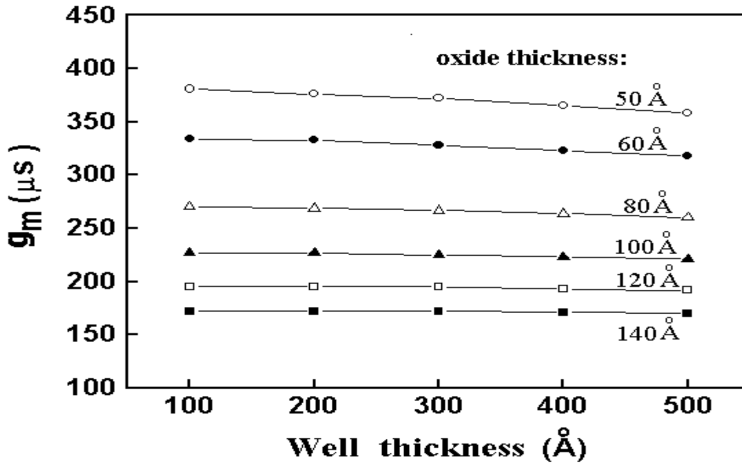



**FIGURE 6.9**
Peak transconductance ($g_m$) as a function of SiGe-well thickness (W/L = 10$\mu$m/1$\mu$m). After S. K. Samanta et al., *Proc. IWPSD-2003*, 2003(540-542).

SiGe channel dominates the device electrical characteristics. Fig. 6.11 shows the peak transconductance vs. Si-cap layer thickness with different oxide thicknesses. The results were obtained using Silvaco-ATLAS device simulator. This figure reveals that the transconductance decreases with increasing Si-cap layer thickness as a consequence of the reduced hole population in the alloy layer.

To maximize the gate-to-channel capacitance and hence to increase the SiGe p-MOSFET transconductance, it is important to minimize both the thickness of the Si-cap and the gate oxide. For a Si-cap thickness below 30 Å, the $p^+$-gate design becomes an attractive alternative to the $n^+$-gate SiGe p-MOSFETs. However, such a thin Si-cap layer does not permit a thermal reoxidation of the source and drain areas after the polySi gates are etched. In addition, since the current flows less than 30 Å away from the gate oxide, interface scattering will degrade the hole mobility. Conversely, with a thin Si-cap, the holes in the SiGe-channel flow closer to the gate and the channel-to-gate capacitance increases. The Si-cap thickness needed to optimize the transconductance is therefore determined by a mobility/capacitance trade-off. The potential improvements in device performance, in terms of carrier confinement and gate-to-channel capacitance, make SiGe transistors with a thin Si-cap very attractive for short-channel applications.

In the following, we present the results of simulations using MOSFET structure with $\delta$-doping spike. The hole distribution in the Si-cap and SiGe-well for structures with and without $\delta$-doping spike is shown in Fig. 6.12. The potential at strong inversion is loosely defined by the condition, $\psi_s = 2\psi_b$,

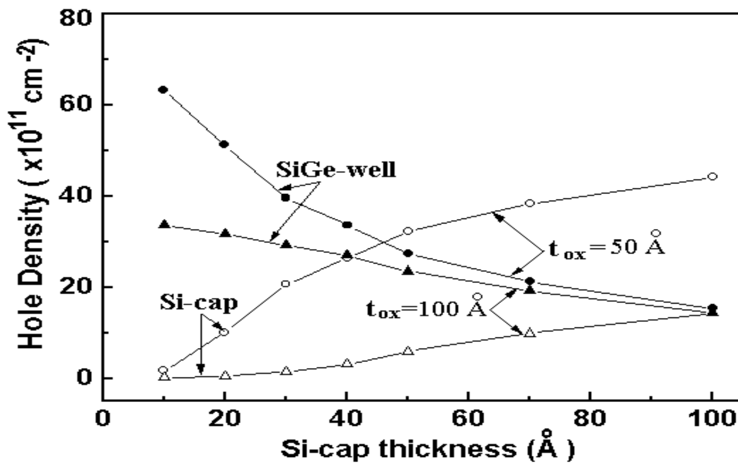



**FIGURE 6.10**
Simulated hole distribution in the buried channel (SiGe-well) and surface channel (Si-cap) as a function of Si-cap layer thickness at a gate voltage ($V_g$ = -2 V). After S. K. Samanta et al., *Proc. IWPSD-2003*, 2003(540-542).

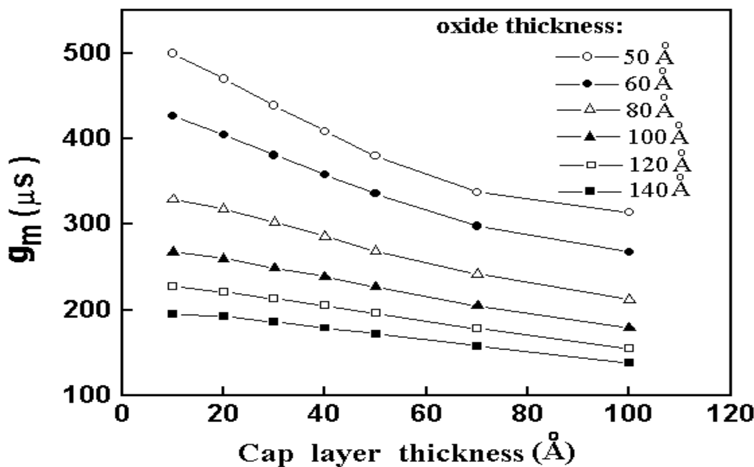



**FIGURE 6.11**
Simulated peak transconductance ($g_m$) as a function of Si-cap layer thickness (W/L = 10$\mu$m/1$\mu$m). After S. K. Samanta et al., *Proc. IWPSD-2003*, 2003(540-542).

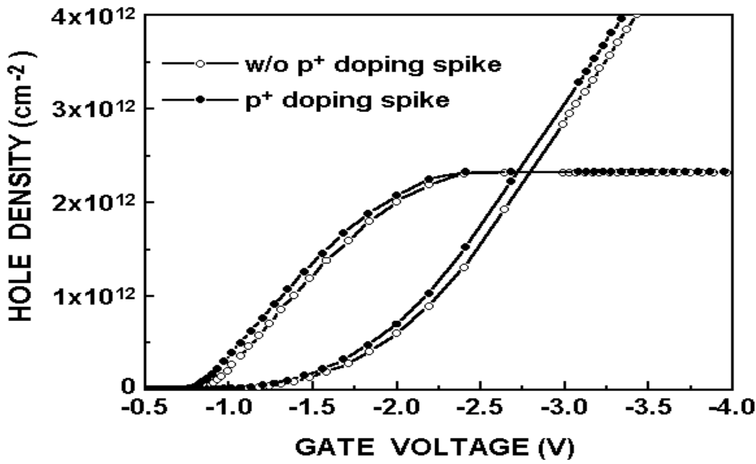



**FIGURE 6.12**
Comparison of hole densities in channel with and without $\delta$-doping spike. After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

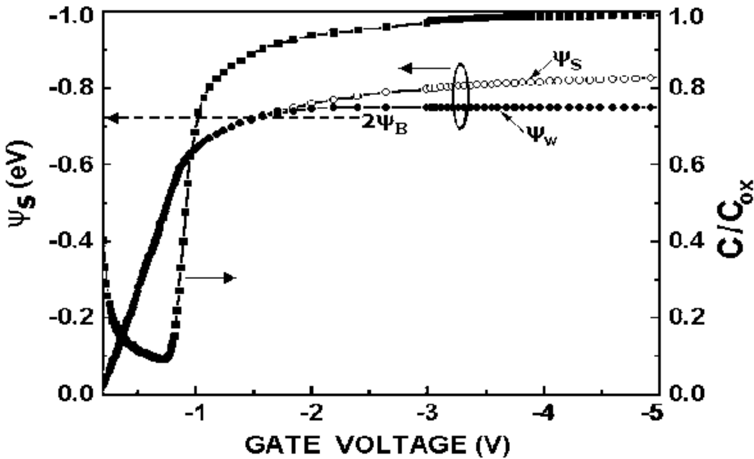



**FIGURE 6.13**
Plot of electrostatic potentials and normalized $C/C_{ox}$. After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

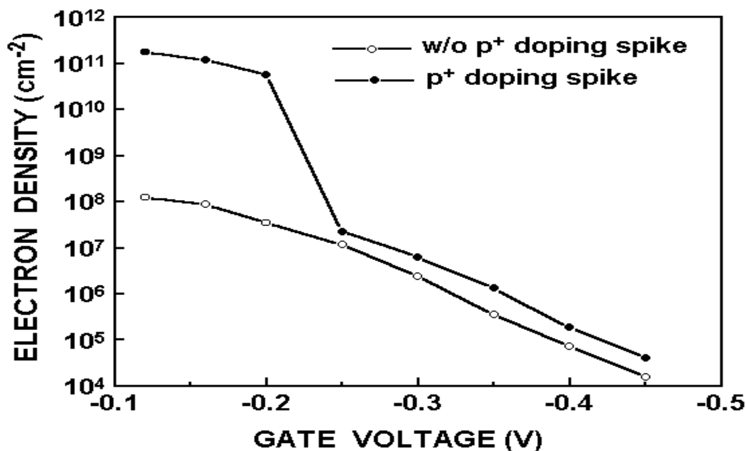



**FIGURE 6.14**
Areal electron density from analytical model. After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

where $\psi_b$ is bulk Fermi potential in the Si-cap. The saturation of carriers in the well is due to potential shielding by the 2DHG at the surface. For the case where $\psi_s$ exceeds $2\psi_b$, $Q_{cap}$ is approximately exponentially dependent on $\psi_s$. Hence an incremental increase in $\psi_s$ by $\phi_t$ (thermal voltage) results in a factor of $e$ increase in surface inversion layer charges [39].

Consequently, any small increase in $\psi_s$ is offset by a commensurate increase in $Q_{cap}$, causing the well potential, $\psi_w$, to remain essentially constant. This explanation for carrier saturation in the well is borne out by Fig. 6.13, which plots the electrostatic potentials, $\psi_s$ and $\psi_w$, as a function of gate bias without $\delta$-doping spike. The corresponding normalized capacitance is shown in Fig. 6.13. Although the hole distribution curves for both with and without $\delta$-doping spike appear similar from Fig. 6.12, the difference lies in the areal electron density in the Si-cap at low gate biases as shown in Fig. 6.14. As expected, the structure with the $p^+$ threshold adjust ($\delta$-doping spike) has a significantly higher density of electrons near zero gate bias, leading to the conclusion that a $\delta$-doping spike would serve to improve the subthreshold characteristic and hole mobility of SiGe-channel p-MOSFETs.

Fig. 6.15 shows the simulated DC output characteristics of $Si_{1-x}Ge_x$ (with x = 0.3) devices for both with and without $\delta$-doping spike at $V_g$ = -2.0, -3.0 and -4.0 V. The drain current for structure with $\delta$-doping spike increases compared to a structure without $\delta$-doping spike. The saturation drain current of $\delta$-doped device is found to be -3.75 mA for 30% Ge at a drain voltage of -2.0 V and gate-to-source voltage of -4.0 V. The drain current is about 141%

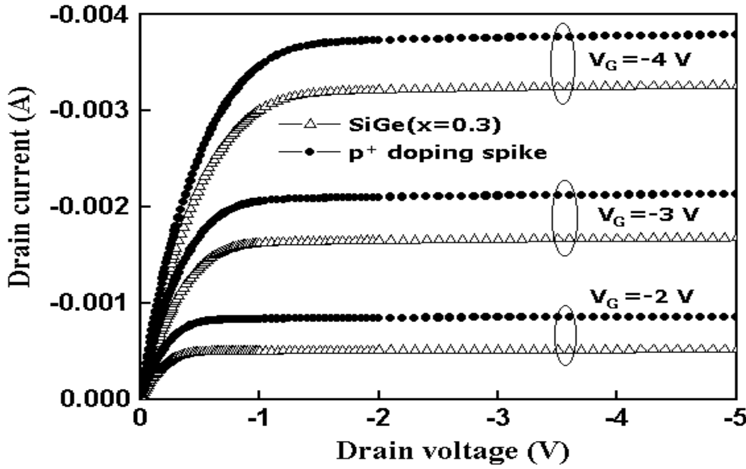



**FIGURE 6.15**
Simulated output characteristics of $Si_{1-x}Ge_x$ devices with different Ge concentration (W/L = 10$\mu$m/1$\mu$m). After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

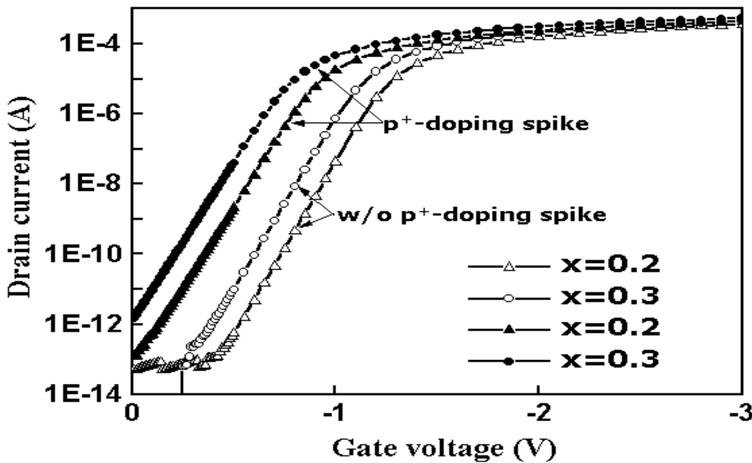



**FIGURE 6.16**
Subthreshold characteristics of $Si_{1-x}Ge_x$ devices at $V_d$ = -0.1 V (W/L = 10$\mu$m/1$\mu$m). After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

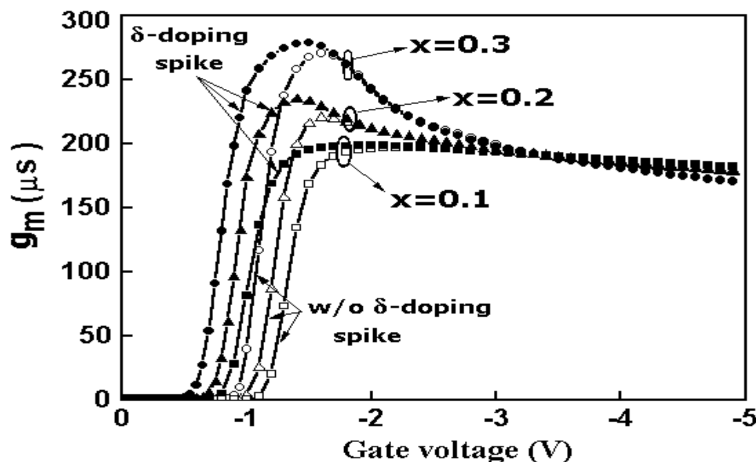



**FIGURE 6.17**
Simulated transconductance characteristics of SiGe p-MOSFETs with comparison between $\delta$-doping spike and without $\delta$-doping spike at $V_d$ = -0.1 V (W/L = 10$\mu$m/1$\mu$m). After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

higher compared to Si devices. The $Si_{0.7}Ge_{0.3}$ transistor with $\delta$-doping spike exhibits a higher (33%) drive current at the same $V_g$ (= -3 V) than control $Si_{0.7}Ge_{0.3}$ devices.

Subthreshold characteristics of $Si_{1-x}Ge_x$ (x = 0.2 and 0.3) p-MOSFETs (with and without $\delta$-doping spike) are shown in Fig. 6.16. Subthreshold swings are found to be 123, 112, and 112 mV/dec (102, 104, and 104 mV/dec) for 10%, 20%, and 30% Ge, respectively, with and without $\delta$-doping spike. The corresponding value is 84 mV/decade for control-Si device. Subthreshold slope for SiGe p-MOSFETs is higher compared to control-Si devices [40]. This is due to the fact that the buried channel is less sensitive to gate control under weak inversion and possibly due to the higher permittivity of $Si_{1-x}Ge_x$ layers [45].

Fig. 6.17 shows the transconductance of $Si_{1-x}Ge_x$ devices computed using silvaco-ATLAS simulator at a low field ($V_d$ = -0.1 V). In simulation, a structure similar to one described in Fig. 6.2 is used with a Si-cap thickness of 50 Å and Ge mole fraction of 10, 20, and 30%. It is found that an increase in Ge fraction increases the $g_m$, when one compares two devices with and without a $\delta$-doping spike. Fig. 6.18 shows the variation of threshold voltage and peak transconductance with Ge concentration. The threshold voltage decreases and peak transconductance increases with increase in Ge concentration for both the devices. The threshold voltage decreases rapidly in $\delta$-doping spike

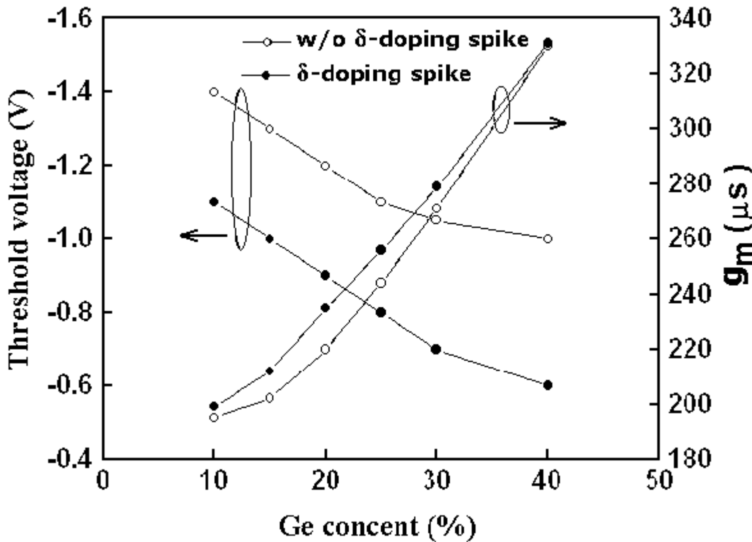



**FIGURE 6.18**
Simulated threshold voltage and peak transconductance as a function of Ge concentration. After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

compared to that without $\delta$-doping. From the plot, it is seen that the presence of a $\delta$-doped layer is advantageous mainly in terms of threshold voltage adjustment with slight improvement in peak transconductance.

## 6.3 SiGe p-MOSFETs on SOI

International Technology Roadmap for Semiconductors (ITRS) [46] identifies the growing challenges for CMOS scaling and elaborate on options being explored to meet future transistor requirements. Silicon-on-insulator has finally been put on the microelectronics technology roadmap because of its high density and improved performance in circuit speed demonstrated. At the same time, SiGe alloys have shown their potential for novel high-performance heterostructure devices and circuits including MOSFETs [47]. There is a great deal of interest in combining two technologies, viz., the SOI-CMOS and the recently developed SiGe technologies. Among the various SOI and SOS technologies, SIMOX is probably the best candidate for the improvement of mo-

**TABLE 6.4**
Reported results of Si-channel p-MOSFETs on SOI, SOS, and SIMOX substrates. After S. Maikap, *Ultrathin Dielectrics on SiGe/SiGeC Layers for Heterostructure MOSFET Applications*, Ph.D. Thesis, Indian Institute of Technology, Kharagpur, 2001.

| Parameter | [7] | [9] | [10] |
|---|---|---|---|
| Channel length, $L_{eff}$ ($\mu$m) | 10 | 1.0 | 0.05 |
| Oxide thickness, $t_{ox}$ (Å) | 65 | 80 | 20 |
| Cap thickness, $t_{cap}$ (Å) | 100 | 70 | 40 |
| Well thickness, $t_{well}$ (Å) | 100 | 120 | 150 |
| Ge concentration, x | 0.3 | 0.2 | 0.0-0.3 |
| Spacer thickness, $t_{sp}$ (Å) | 1600 | 100 | - |
| Channel doping, $N_d$ ($cm^{-3}$) | $1\times10^{15}$ | $4\times10^{17}$ | undoped |
| Buried oxide thickness (Å) | 4100 | 1000 | 4500 |
| $g_m$ at 300 K (mS/mm) | 0.72 | $\sim$ 3.5 | - |
| Improved drive current (%) | - | - | 70 |
| Substrate type | SIMOX | SOS | SOI |

bility in SiGe-channel p-MOSFETs. Nayak et al. [7] reported SiGe-channel p-MOSFET on SIMOX substrates. The device consists of a $Si/Si_{0.7}Ge_{0.3}/Si$ channel, which is grown pseudomorphically on a SIMOX substrate. The effective channel mobility (181 $cm^2$/V-s) of this device was found to be 90% higher than that of an identically processed conventional SIMOX device. Table 6.4 shows some reported device parameters for SiGe-channel p-MOSFETs on SOI, SOS, and SIMOX substrates.

In the following, we present a comprehensive study on the scaling properties of SiGe-channel p-MOSFETs on Si (SiGe-bulk) and p-MOSFETs on SOI (SiGe-SIMOX) by using numerical simulation. A 1D Poisson solver and Silvaco-ATLAS are used. A SiGe-SIMOX device is similar to a SiGe-channel p-MOSFET on Si as shown in Fig. 6.1, with the main distinction being the presence of a buried oxide (thickness: 410 nm) and a spacer layer (thickness: 160 nm) in the former. One way to verify the hole confinement in the quantum well is to study the low frequency capacitance where a plateau in the inversion capacitance signifies the hole confinement.

Fig. 6.19 shows the computed low-frequency C-V characteristics of SiGe-SIMOX and SiGe-bulk structures for both the $n^+$- and $p^+$-polySi gate using the 1D Poisson solver. The plateau in the curve, between gate bias of about 0 to -1.0 V for $p^+$-polySi and -1.0 to -2.0 V for $n^+$-polySi gate, is the result of hole confinement in the SiGe-well. The plateau in accumulation region for the SiGe-SIMOX structure extends to wider range of gate voltage when compared to that for a SiGe-bulk structure. The band diagram is shown in Fig. 6.20.

In Fig. 6.21, the simulated hole distribution as a function of band bending in both SiGe-SIMOX and SiGe-bulk devices is shown. For the SiGe-bulk device,

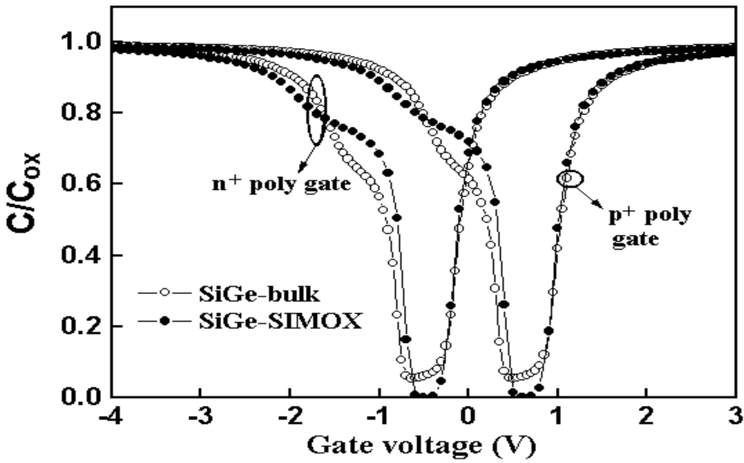



**FIGURE 6.19**
Low-frequency capacitance-voltage (C-V) characteristics of SiGe-bulk and SiGe-SIMOX devices with x = 0.3 using 1D Poisson solver. After C. K. Maiti et al., *Proc. IWPSD-2003*, 2003(543-545).

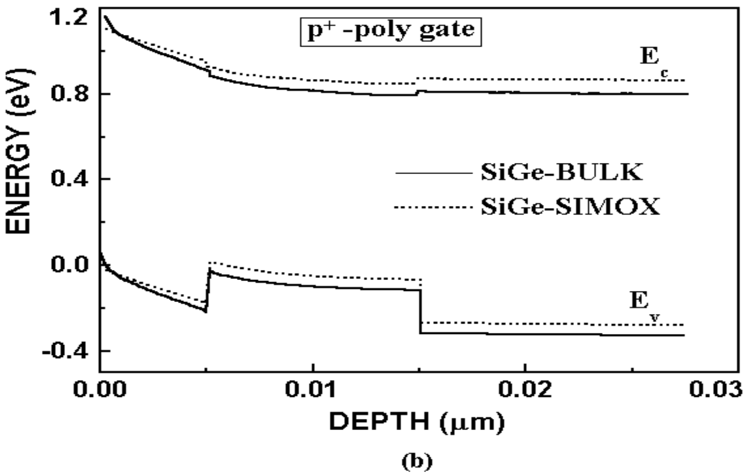



**FIGURE 6.20**
Band diagram (p$^+$-polySi gate) of SiGe-bulk and SiGe-SIMOX devices with x = 0.3 using 1D Poisson solver. After C. K. Maiti et al., *Proc. IWPSD-2003*, 2003(543-545).

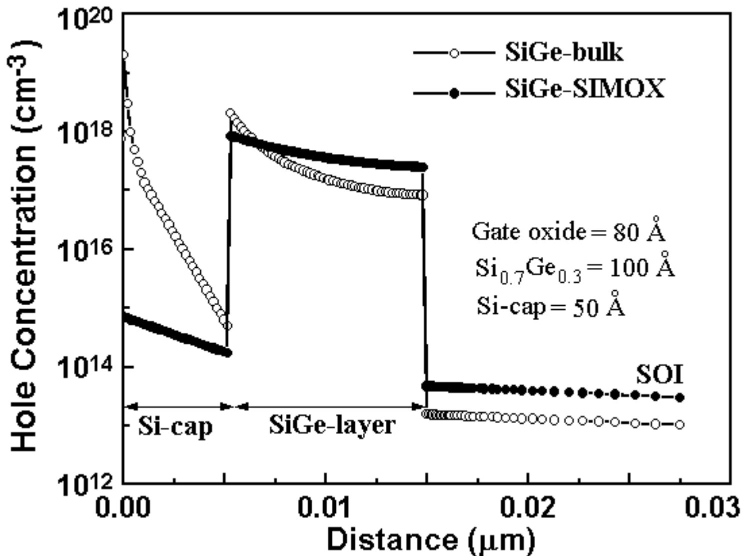



**FIGURE 6.21**
Hole distribution in SiGe-bulk and SiGe-SIMOX. After C. K. Maiti et al., *Proc. IWPSD-2003*, 2003(543-545).

the same well and gate oxide (as in SiGe-SIMOX) parameters are used, and the substrate doping is taken to be $5\times10^{15}$ $cm^{-3}$. The hole concentration at the Si surface for the SiGe-SIMOX is about two orders of magnitude lower than that for the SiGe-bulk device, which means that the channel conduction through the parasitic surface channel (at $Si/SiO_2$ interface) is significantly diminished in SiGe-SIMOX due to reduced band bending at the surface. The reduced band bending also results in a more uniform hole concentration in the SiGe-well for the SIMOX device. For the same Si cap layer thickness, reduced $Si/SiO_2$ surface scattering for the SiGe-SIMOX device results in a further improvement in the channel mobility.

A large valence band discontinuity at the top $Si/Si_{1-x}Ge_x$ heterointerface is essential to increase the hole population in the SiGe channel and ensure adequate confinement up to a high gate voltage. The improved confinement with increased Ge concentration (SiGe-SIMOX and SiGe-bulk devices) is illustrated in Fig. 6.22, where the $n^+$-polySi gate MOSFET crossover voltage is plotted. A higher cross-over voltage is observed for SiGe-SIMOX devices compared to SiGe-bulk devices. The ratio of holes in the Si-cap to the SiGe channel for two different gate voltages ($V_g$ = -1.5 V and -2.5 V) are plotted in Fig. 6.22. As expected, more carriers flow in the Si-cap layer as the gate voltage rises. In order to ensure that the majority of holes flow in the SiGe-channel for gate voltage equal to or more negative than -2.5 V, the Ge

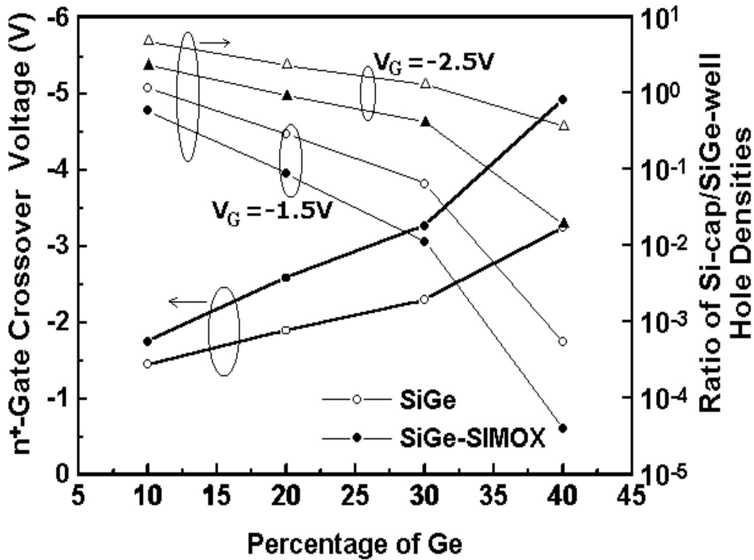



**FIGURE 6.22**
$n^+$-gate cross-over voltage vs. Ge mole fraction for SiGe-bulk and SiGe-SIMOX devices. After C. K. Maiti et al., *Proc. IWPSD-2003*, 2003(543-545).

concentration must rise above 30%. The ratio of holes in the Si-cap over those in SiGe-channel are observed to be lower for SiGe-SIMOX devices compared to SiGe-bulk devices.

Fig. 6.23 compares the I-V characteristics obtained using Silvaco-ATLAS 2D device simulator for SIMOX devices made on Si and SiGe substrates at room temperature. All devices show both linear and saturation regions. The drain current of the SiGe-SIMOX device is 222% higher, for voltages equal to or higher than -3.0 V and the Ge percentage must be above 30%, compared to that for the Si-SIMOX device. This shows that the circuits built using the SiGe-SIMOX device will have potentially high speed due to the higher drive current. The turn-on and turn-off characteristics of a device is largely determined by the current drive as well as the peak transconductance of the device. The improvement in the drain current of the SiGe-SIMOX device over the SIMOX device is significant.

A plot of linear transconductance ($g_m$) vs. gate voltage ($V_g$) for a 1 $\mu$m channel length device at a low field ($V_d$ = -0.1 V) is shown in Fig. 6.24. The improvement in peak $g_m$ for SiGe-SIMOX device over that of Si-SIMOX is evident from this figure. The $g_m$ of SiGe-SIMOX is found to increase with an increase in Ge mole fraction and decreases rapidly at high transverse field. The shape of the transconductance curve for the SiGe-SIMOX is slightly different than that of the Si-SIMOX device. The transconductance of the SIMOX

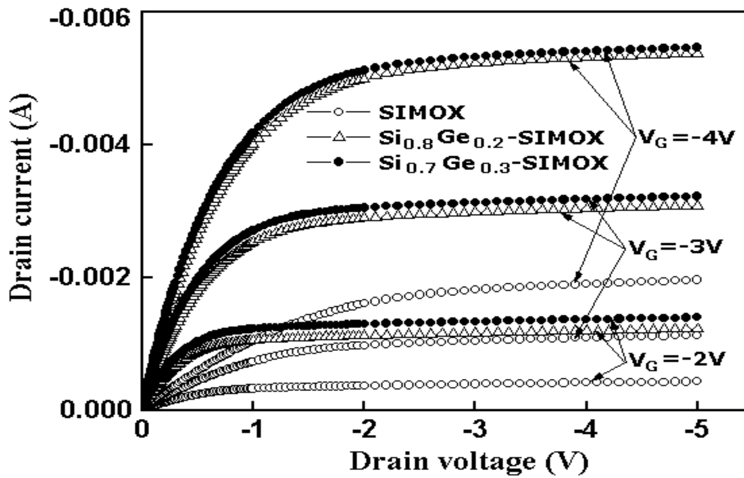



**FIGURE 6.23**
Output characteristics for SiGe-SIMOX and SIMOX. After C. K. Maiti et al., *Proc. IWPSD-2003*, 2003(543-545).

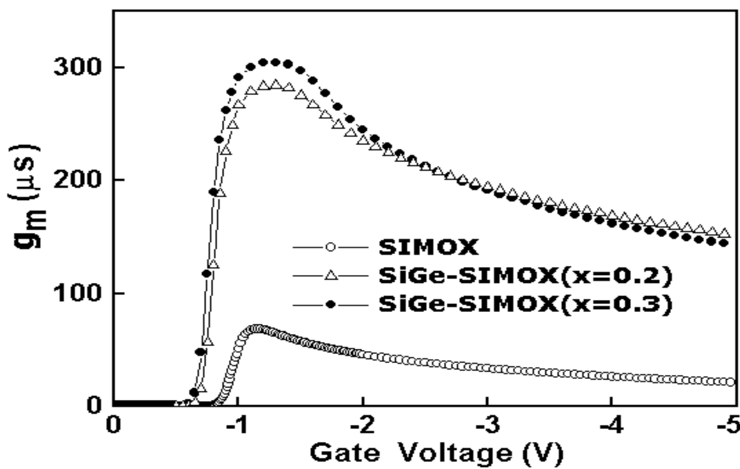



**FIGURE 6.24**
Comparison of linear transconductances of SiGe-SIMOX and Si-SIMOX devices (W/L = 10 $\mu$m/1 $\mu$m). After C. K. Maiti et al., *Proc. IWPSD-2003*, 2003(543-545).

device exhibits a well-defined peak near threshold, similar to a bulk-Si device. In case of SiGe-SIMOX, three regions are observed in the transconductance characteristics:

(a) at low gate voltages (close to the threshold voltage, $V_g \approx$-1.0V), the effective capacitance is small (series combination of oxide capacitance and Si-cap layer capacitance) but the peak $g_m$ is high because the holes are mainly confined to the well. Therefore, the peak $g_m$ as well as the channel mobility, will be high.

(b) The peak $g_m$ value remains nearly same for a range of gate voltages ($V_g$ = -1.0 to -1.5V). As the gate voltage is above -1.0V, carriers start to build up at the surface channel, and for these carriers the channel mobility is low due to increased $Si/SiO_2$ surface scattering but the effective gate capacitance is high due to the ultrathin oxide (80 Å). Hence, the net effective transconductance still remains high.

(c) As $V_g$ exceeds -1.5V, the peak transconductance falls rapidly, with the carriers in the surface channel starting to dominate the total channel conduction. As the effective field becomes more negative, the surface channel turns on and the transconductance decreases rapidly. This type of transconductance behavior has been observed earlier for SiGe p-MOSFETs on SOS [9].

## 6.4 SiGeC Hetero-FETs

The growth of pseudomorphic strained-$Si_{1-x}Ge_x$ alloys on Si has generated considerable interest because of their potential application in high performance p-MOSFETs due to their compatibility with Si technology [47]. But the use of strained-$Si_{1-x}Ge_x$ alloys has some practical limitations in terms of the critical layer thickness and thermal budget for device processing. By incorporating smaller-sized C atoms in the $Si_{1-x}Ge_x$ alloy, the strain can be compensated [48, 49], allowing additional flexibility in strain engineering along with the improvement of the thermal stability of the resultant $Si_{1-x-y}Ge_xC_y$ alloy layers. Since C is smaller than both Si and Ge, it can introduce local tensile strain and compensate the compressive strain introduced by Ge, leading to complete strain compensation at a Ge-to-C ratio of 8:1. $Si_{1-x-y}Ge_xC_y$ layers are more robust in terms of the allowable thermal budget during processing. In a type-I $Si/Si_{1-x-y}Ge_xC_y$ heterostructure band alignment, a ternary alloy shows a higher valence band offset as compared to that of SiGe for the same lattice mismatch to Si. This makes the ternary layer attractive for enhanced hole mobility p-MOSFET devices with a Si spacer between the channel and the gate oxide to prevent the pile of unoxidized Ge and C behind the oxide front.

The band alignment in the heterostructure favors the confinement of holes

in the valence band well leading to dominant carrier transport in the high-mobility channel layer, with reduced scattering from surface roughness and oxide charges. The device performance can be optimized by choosing an appropriate thickness of the unconsumed Si cap on top of the channel layer in this buried $Si_{1-x-y}Ge_xC_y$ pMOS device. Several authors have demonstrated the enhancement of hole mobility in buried-channel partially strain-compensated SiGeC p-MOSFET device over that of the control-Si devices [42, 50].

Partially strain-compensated $Si_{0.793}Ge_{0.2}C_{0.007}$ (C compensating approximately the strain equivalent to 8-10% Ge) epitaxial layer was grown by ultra-high vacuum chemical vapor deposition on an n-Si substrate at a temperature of 550$^o$C. The growth sequence starts with a 700 Å silicon buffer layer, followed by the growth of 400 Å in situ phosphorus-doped $Si_{0.793}Ge_{0.2}C_{0.007}$ and a starting Si cap with a thickness of 70 Å. The sample had a channel doping of 1-2$\times10^{17}$ atoms cm$^{-3}$ as obtained from secondary ion mass spectroscopy (SIMS) analysis. Rapid thermal oxidation (RTO) at 950$^o$C for 3 min in an oxygen ambient was used to grow a high-quality gate oxide of a thickness of 80 Å, leaving an unconsumed Si cap layer of 30-35 Å. A self-aligned n$^+$-poly-gate process was used to fabricate p-MOSFET devices with a fixed width of 10 $\mu$m and a channel length varying from 0.8-10 $\mu$m. The p$^+$ source/drain (S/D) regions were formed by $BF^+{}_2$ implantation at 20 keV with a dose of 2$\times10^{17}$ cm$^{-2}$. The dopant activation was carried out at 950$^o$C for 1 min by rapid thermal annealing (RTA). Contacts to the S/D regions were formed by sputtering Al (1% Si) followed by a post-metal anneal at 450$^o$C for 30 min in forming gas. Epitaxial Si p-MOSFETs, as the control devices, were also fabricated in the same run for comparison. The characteristics of the MOSFET device were measured over a wide range of temperatures and electric fields.

Fig. 6.25 shows the subthreshold ($I_d$-$V_g$) characteristics of $Si_{0.793}Ge_{0.2}C_{0.007}$ devices with a channel length of 10 $\mu$m at 300 K and 77 K. The subthreshold slopes are found to be 94 mV/decade at 300 K and 42 mV/decade at 77 K for $Si_{0.793}Ge_{0.2}C_{0.007}$ devices. In comparison, the corresponding values are found to be 81 and 25 mV/decade for the control Si devices. The subthreshold slopes for ternary devices are higher compared to the control Si devices owing to the higher heterointerface states in the alloy layers. The output characteristics of SiGeC and the control Si devices at room temperature are compared in the inset of Fig. 6.25. Both the devices exhibit good saturation and turn-off characteristics. At the same gate voltage, the $Si_{0.793}Ge_{0.2}C_{0.007}$ MOSFET shows a higher drive current than the control Si device.

Fig. 6.26 shows the computed hole densities in the buried $Si_{0.793}Ge_{0.2}C_{0.007}$ and the parasitic Si cap as a function of gate voltage at room temperature and 77 K. The threshold voltages of the buried channel from the hole density are found to be -1.66 V and -1.9 V at 300 K and 77 K, respectively. These values are in good agreement with the experimentally obtained values of -1.64 V and -1.93 V, respectively. The crossover voltage, defined as the transition to a predominantly surface-channel device from a buried-channel device, is found to be -3.9 V at both temperatures. Fig. 6.26 clearly shows the charge

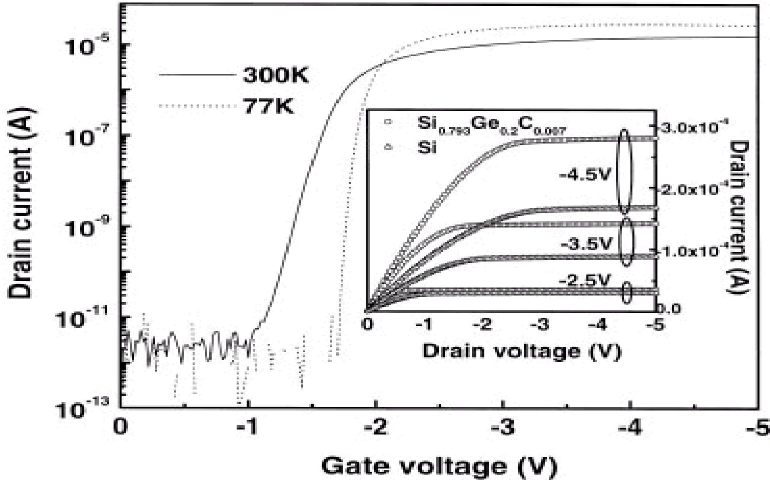



**FIGURE 6.25**

$I_d$-$V_g$ characteristics of $Si_{0.793}Ge_{0.2}C_{0.007}$ device at room temperature and 77 K. Measured room temperature output characteristics of $Si_{0.793}Ge_{0.2}C_{0.007}$ and control Si p-MOSFET devices are shown in the inset. After G. S. Kar et al., *Semicond. Sci. Technol.*, Vol. 17, 2002(471-475).

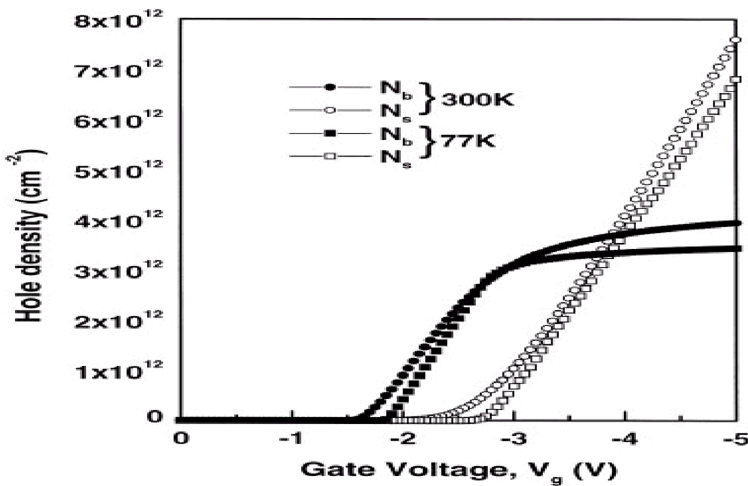



**FIGURE 6.26**

Variation of hole concentration in surface ($N_s$)and buried ($N_b$) channels as a function of gate bias at room temperature and 77 K. After G. S. Kar et al., *Semicond. Sci. Technol.*, Vol. 17, 2002(471-475).

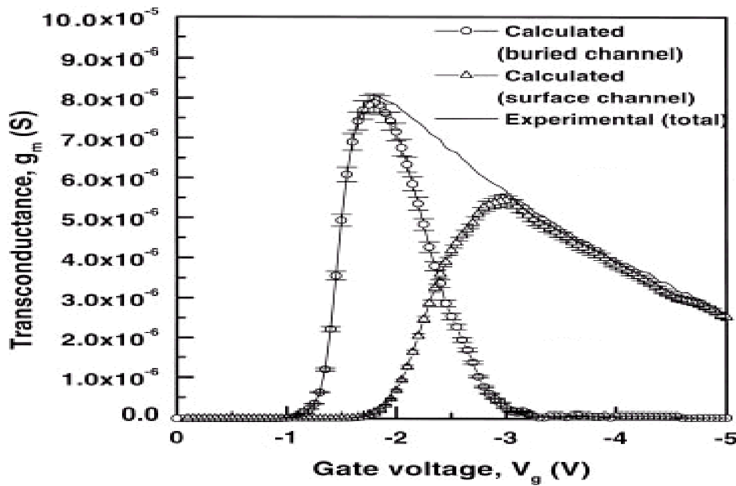



**FIGURE 6.27**

Calculated transconductance in surface and buried channels, along with experimentally obtained total transconductance (solid) plots at room temperature. After G. S. Kar et al., *Semicond. Sci. Technol.*, Vol. 17, 2002(471-475).

screening effect associated with a buried-channel structure. The saturation of the buried channel current can be explained as follows. Once a large number of holes flow in the Si cap layer, the charge screens the gate potential and fewer holes are added to the SiGeC channel with increasing gate voltage, limiting the maximum current contributed by high-mobility holes.

Fig. 6.27 shows the calculated linear transconductance ($g_m$) (open up triangle and circle) characteristics for both surface and buried channels at room temperature, along with the experimentally measured total transconductance of the device. A sharp fall in the calculated $g_m$ value in the buried channel compared to that of the Si surface channel reveals the saturation of holes in the SiGeC channel at a higher gate bias. The peak transconductance in the buried ternary alloy is 1.48 times higher than that in the cap Si. The effective mobility of holes in the ternary alloy has been calculated from the net current in the buried layer using the simulated hole density of Fig. 6.26. Fig. 6.28 shows the variation of extracted effective mobility in the $Si_{0.793}Ge_{0.2}C_{0.007}$ channel as a function of effective field at room temperature, and 77 K. $Si_{0.8}Ge_{0.2}$ (layer with no carbon) devices data are also shown for comparison.

As it is difficult to form high-quality thermal gate oxide on $Si_{1-x}Ge_x$, an alternate approach is to use a deposited gate oxide, such as a high dielectric constant gate dielectric. This avoids the need for the deposition of the Si cap and thereby rendering a surface channel device [51, 52]. This is particularly attractive for a $Si_{1-x}Ge_x$ or pure Ge channel because it avoids the problems of Ge segregation during high thermal budget gate oxidation. As

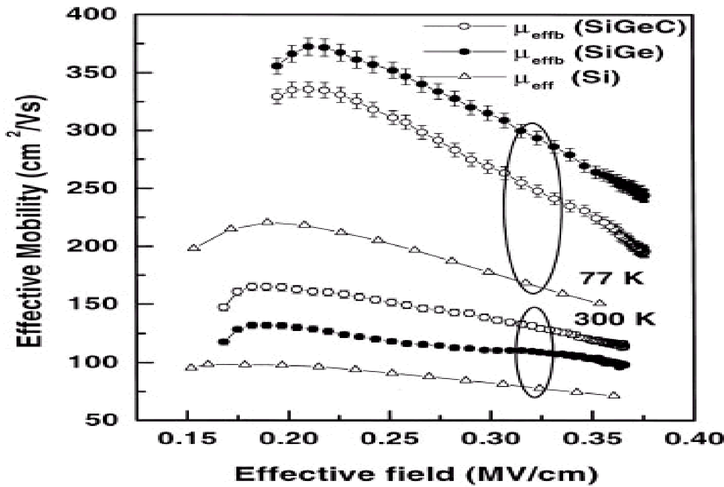



**FIGURE 6.28**
Variation of effective mobility with effective field in the buried $Si_{0.793}Ge_{0.2}C_{0.007}$ channel, along with the buried $Si_{0.8}Ge_{0.2}$ channel and control Si device at room temperature and 77 K. After G. S. Kar et al., *Semicond. Sci. Technol.*, Vol. 17, 2002(471-475).

compared to Si-$SiO_2$ p-MOSFETs, Si high-k p-MOSFETs in general exhibit a mobility degradation of about 25%. Thus, SiGe-channel devices with high-k as gate dielectric can provide a means to overcome the mobility degradation on Si caused by using high-k materials as gate dielectric. This mobility enhancement enables a higher drive current, even at channel lengths down to 90 nm. Unfortunately, other important transistor parameters such as the subthreshold characteristics show that the drain-induced barrier lowering is comparable for both SiGe and Si control devices, the subthreshold slope and junction leakage are significantly worse for SiGe devices. Also, the significant difference in $V_t$ between the $Si_{0.8}Ge_{0.2}$ and Si p-MOSFETs can be attributed to the bandgap difference between $Si_{0.8}Ge_{0.2}$ and Si. The bandgap difference is manifested mostly in the valence band, and for Ge mole fraction of 0.2 a 168 mV valence band offset is obtained.

## 6.5 SiGe-based HEMTs

The need for advanced RFICs, combined with the recent progress in the growth of pseudomorphic SiGe, have led to increased interest in silicon-based heterostructure MODFETs. The key design issues for short-channel strained-

SiGe heterojunction MODFETs have been addressed in detail [53]. In the following, the high-frequency performance analysis of a 2DEG strained-Si/SiGe HEMT with Schottky gate is performed by computer simulation using the 2D drift diffusion (DD) and energy balance (EB) models. The material parameters for both strained-Si and $Si_{1-x}Ge_x$ have been modeled and are used in the simulation. The simulation results are also verified with published experimental results on strained-Si/SiGe HEMT devices. Ismail et al. [54] have reported n-channel devices having room temperature mobilities of 2600 $cm^2$/V-s using a strained-Si channel grown on a relaxed-SiGe buffer. A high-transconductance n-type si/SiGe modulation-doped field-effect transistor has been created and exhibits 600 mS/mm at 77 K with a gate length of 0.25 $\mu$m.

To improve the rf performance, the optimum epitaxial architecture of a fully pseudomorphic strained-Si/SiGe MODFET heterostructure, viz., position and level of the doping at the front side, backside, and even inside of the 2DEG Si-channel are examined. The effect of thickness of various layers on the performance of HEMTs with gate lengths down to 0.18 $\mu$m are also investigated. The strained-Si channel combined with a constant Ge profile in the buffer layer was used as the input of the ATLAS device simulator. The model parameters are incorporated through c-interpreter function for better flexibility. The simulation results are optimized with experimental data available to validate the material parameters being used. The purpose is to compare different Si-based HFET geometries and to identify which are most suitable for high-frequency applications other than logic. The results of the present study serve to illustrate the extent to which III-V technology can be supplanted by Si-based devices.

In simulation, the gate was approximated as a metal contact having a work function of 5.8 eV, similar to the work function of Pt/Au. All devices are biased with a source-drain voltage of 2.5 V. This bias condition was chosen as representative of the voltage levels suitable for use in low-power mobile systems where SiGe microwave devices are likely to find applications.

The heterostructure device schematic considered for simulation is shown in Fig. 6.29. The basic device structure is the same as that of a reported experimental device [55], grown by molecular beam epitaxy (MBE). The layer sequence consists of a relaxed graded-SiGe buffer on a p-Si substrate (100 $\Omega$-cm), an Sb-doped $Si_{0.6}Ge_{0.4}$ backside doping layer (4 nm, $8\times10^{18}$ $cm^{-3}$), an undoped $Si_{0.6}Ge_{0.4}$ spacer (3 nm), the 9-nm thick strained-Si quantum well serving as the 2DEG channel, a $Si_{0.6}Ge_{0.4}$ spacer (93 nm), an Sb-doped supply layer (4 nm, $1.5\times10^{19}$ $cm^{-3}$), an undoped $Si_{0.6}Ge_{0.4}$ cap spacer (10 nm) and a 5-nm Si cap layer. Ohmic source and drain contacts, defined by phosphorous implantation (20 keV, $2\times$ $10^{15}$ $cm^{-3}$), rapid thermal annealing (RTA: $625^oC$, 45s in $N_2$), Ti/Pt/Au deposition and subsequent alloying yield contact resistances around $4\times$ $10^{-7}\omega$ -$cm^2$. Pt/Au T-shaped Schottky gate contacts with footprints of $L_g$ = 180 nm gate length were asymmetrically arranged at a source-gate distance of 200 nm for series resistance reduction. The effective source drain spacing was 1.5 $\mu$m, corresponding to the implanted source and

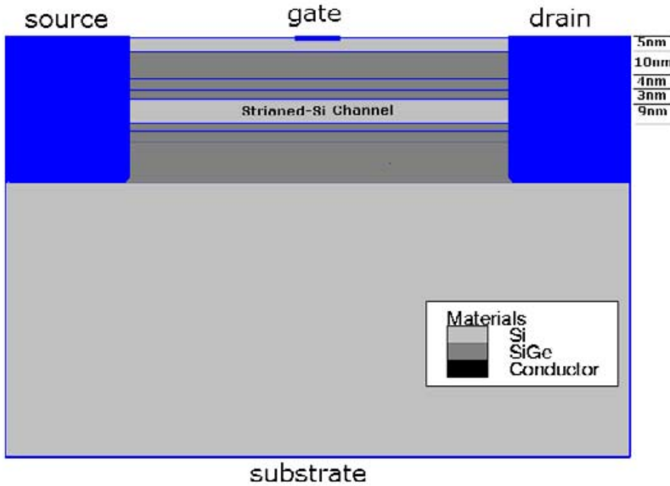



**FIGURE 6.29**
Schematic of a Pt/Au-Schottky gate strained-Si HEMT used in the simulation. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

drain contact zones. The greater source-to-drain separation means that short-channel effects such as punch-through can be suppressed at lower substrate doping levels than in equivalent gate length self-aligned devices. Furthermore, if a Schottky gate is to be made using a nonrefractory metal, such as Pt, then a greater separation of source and drain becomes a necessity. The SiGe cap layer contains the dopant which supplies carriers to the channel. The dopant is set back 3 nm from the heterojunction to prevent ionized impurity scattering in the channel and is contained in a thin layer of thickness 4 nm unless otherwise stated.

Fig. 6.30 shows the simulation and experimental $I_d$-$V_{ds}$ characteristics of a HEMT. The rf characteristics of the devices are shown in Fig. 6.31. The current gain and maximum unilateral gain MUG against frequency curves are derived from S-parameter measurements in the range of 2 to 40 GHz. The simulation of maximum unilateral gain (MUG) is different from experimental results at low frequency because of the extrinsic passive lumped elements of real device. However, a unity current gain cutoff frequency $f_t$ of 46 GHz and a maximum frequency of oscillation $f_{max}$ of 81 GHz were extrapolated at $V_g$ = -0.2V and $V_{ds}$ = 2.5 V are confirmed from the simulation. This high $f_{max}$ for a Si-based hetero-FET is attributed to the low source resistance ($R_s \leq 0.2\omega$/mm), low resistance of the gate ($R_g = 290\,\omega$/mm), the elevated $C_{gs}/C_{gd}$ ratio of $\sim$ 9 to 11 and the high current densities due to the increased carriers

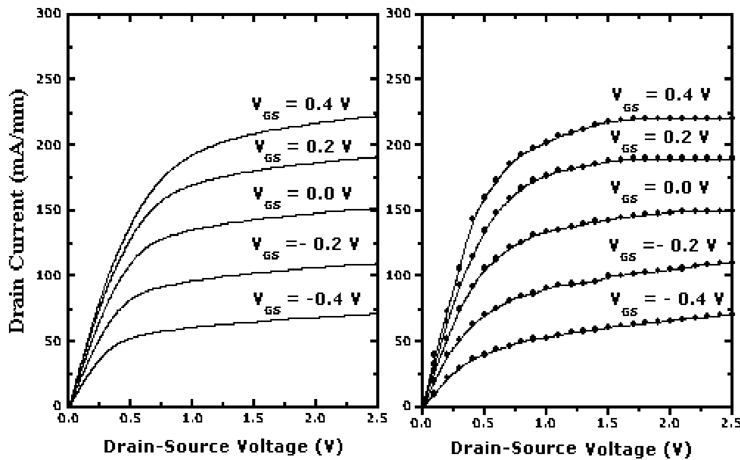



**FIGURE 6.30**
Comparison of simulation with experimental results of DC output characteristics of a 0.18 $\mu$m gate length strained-Si HEMT. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

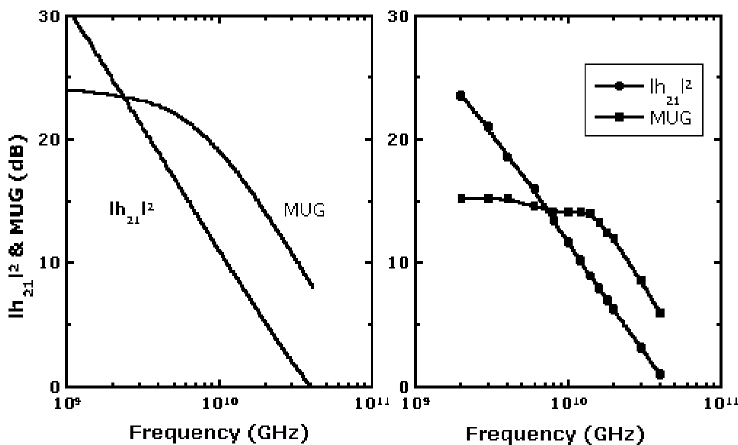



**FIGURE 6.31**
Comparison of simulation with experimental current gain and maximum unilateral gain (MUG) against frequency curves, derived from S-parameter. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

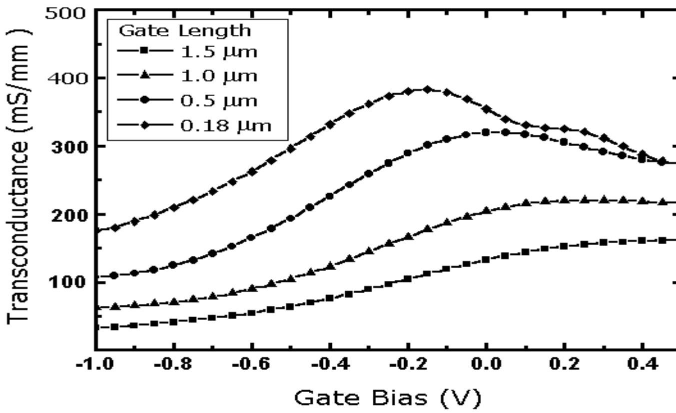



**FIGURE 6.32**
Transconductance of a depletion HEMT with downscale layout at $V_{ds} = 2.5$ V. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

in the $Si_{0.6}Ge_{0.4}$/strained-Si/$Si_{0.6}Ge_{0.4}$ quantum well channel configuration.

## 6.6 Design Issues

In this section, the rf potential of HEMT is examined for gate length down to 0.18 $\mu$m. To improve the rf performance, the position and the doping level at the front and backside and the thickness of the different layers will be considered. The strained-Si channel is maintained at a thickness of 10 nm throughout and is nominally undoped.

### 6.6.1 Gate Engineering

Fig. 6.32 shows the dependence of transconductance on gate length $L_g$ of depletion mode of HEMT. The effect of the gate length variation on $f_t$ and $F_{max}$ is illustrated in Fig. 6.33 for devices having a undoped cap layer of 5 nm and in other respects, the device geometries are identical. Simulation results indicate that the gate is unable to switch the device off at the lowest gate length 0.1 $\mu$m. The increase of $g_m$, with increasing gate voltage is related to the turn-on of the device, while the decrease at higher gate voltage is due to screening of the channel from the gate by the turn-on of a parasitic channel in the Si cap layer. The $f_t$ increases from 46 GHz to 100 GHz, while $g_m$ increased

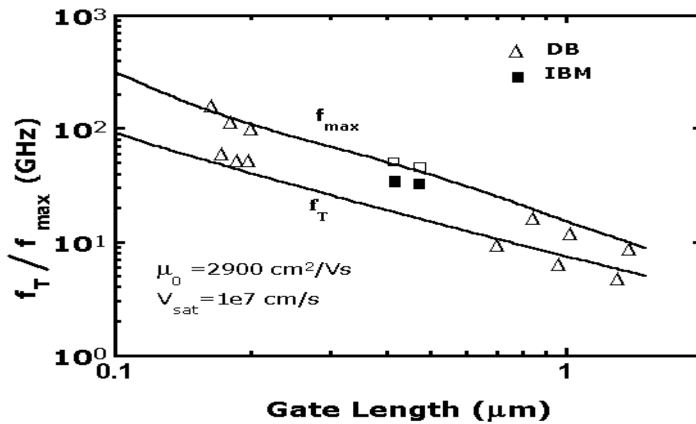



**FIGURE 6.33**
Simulated cutoff frequency and maximum oscillation frequency as a function of gate length at $V_{gs}$ = -0.2 V and $V_{ds}$ = 2.5 V. The data corresponds to the experimental results of DB and IBM. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

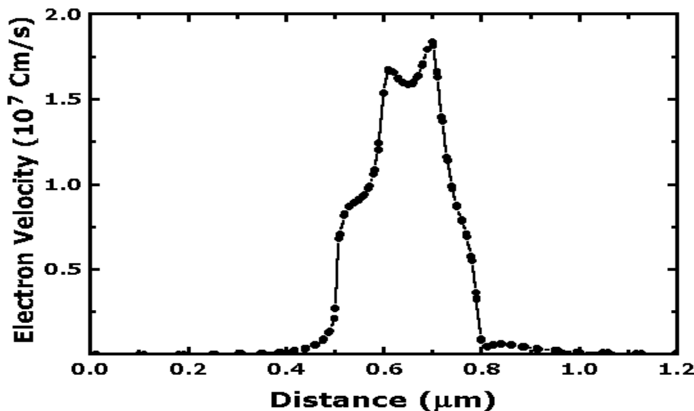



**FIGURE 6.34**
Electron velocity profile for 0.1 $\mu$m gate length HEMT. Velocity overshoot is observed closer to the drain. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

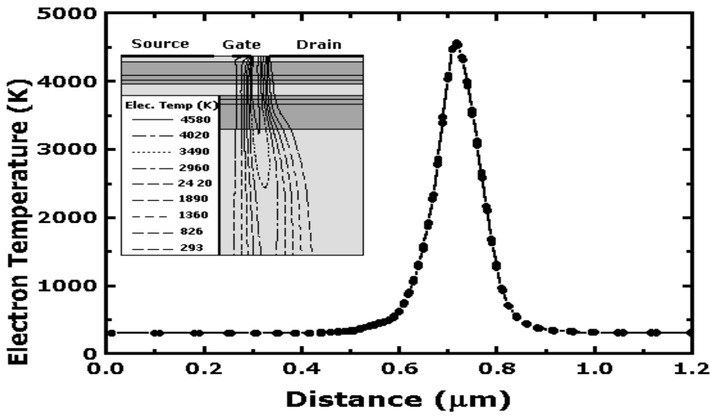



**FIGURE 6.35**
Electron temperature distribution in the strained-Si channel HEMT at $V_{gs}$ = -0.2 V and $V_{ds}$ = 2.5 V. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

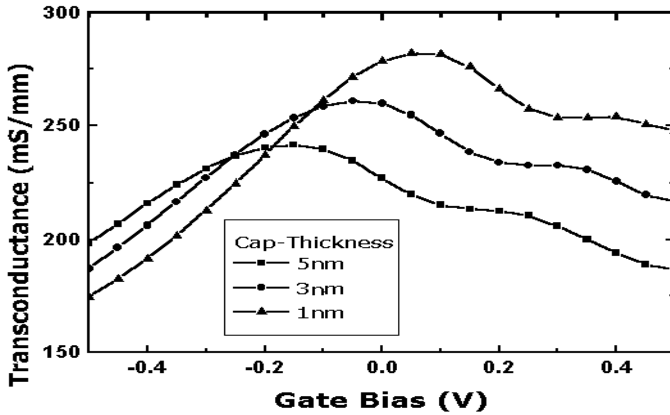



**FIGURE 6.36**
The impact of cap layer thickness on the transconductance. Decrease in the cap layer thickness increases transconductance. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

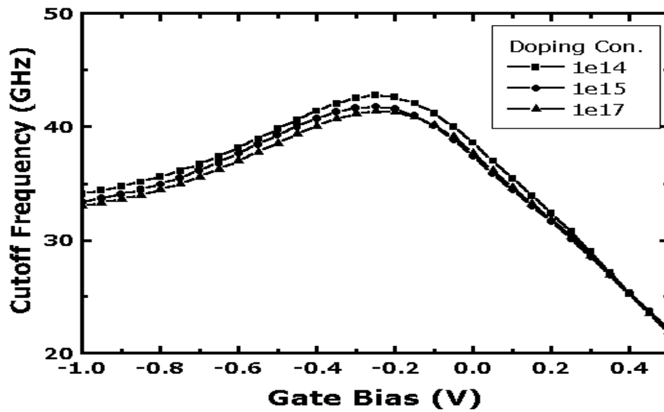



**FIGURE 6.37**
The impact of cap layer doping level on the cutoff frequency at $V_{ds} = 2.5$ V. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

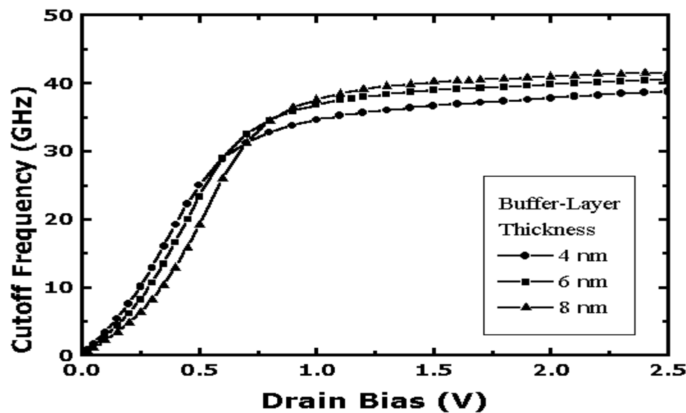



**FIGURE 6.38**
Cutoff frequency vs. drain voltage for different buffer layer thickness for strained-Si HEMT. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

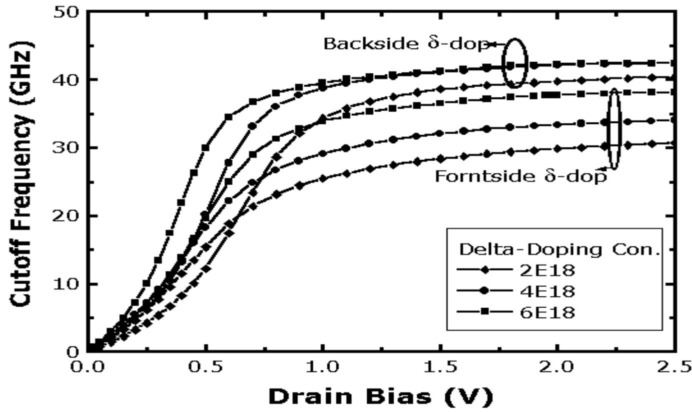



**FIGURE 6.39**
The impact of $\delta$-doped layer on the cutoff frequency for the drain bias $V_{gs}$ = 0.2 V. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

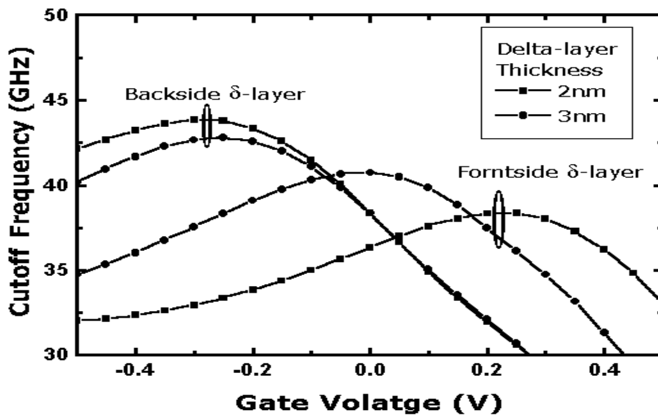



**FIGURE 6.40**
Simulated cutoff frequency as a function of drain voltage for different $\delta$-doped layer thickness of an strained-Si channel HEMT. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

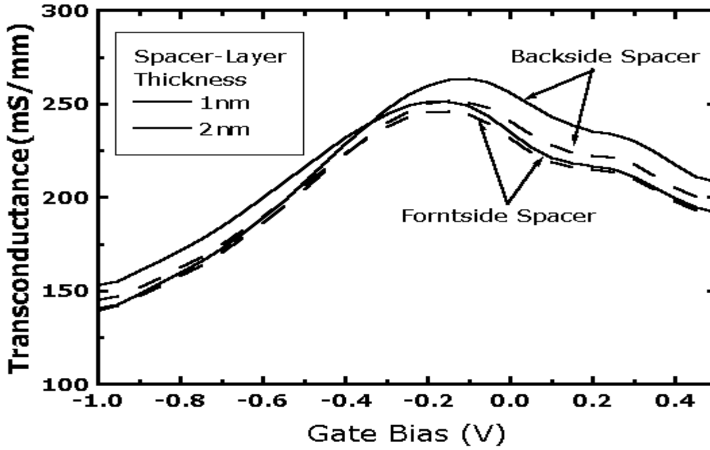



**FIGURE 6.41**
Gate voltage dependence of the transconductance for different spacer-layer thickness. After B. Senapati et al., *Proc. IWPSD-2003*, 2003(833-835).

from 300 mS/mm to 490 mS/mm.

Conventional drift-diffusion (DD) model is not applicable for the simulation of submicron devices. The hydrodynamic model is preferred for determination of velocity overshoot and electron temperature distribution for scaled device layouts. Fig. 6.34 illustrates the longitudinal velocity as a function of position in the quantum well of a 0.1 $\mu$m gate length device. Velocity overshoot is observed close to the drain with a peak velocity of $2\times10^7$ cm/s. The overshoot is important in achieving high transconductance for a device as it introduces a larger current flow along the quantum well. The first peak near the source is due to the structure used. Due to velocity overshoot, the peak electron temperature also occurs at the same location. The electron temperature profile, along with device structure, is shown in Fig. 6.35.

### 6.6.2 Layer Design

Attention is now focused on the SiGe cap layer, in particular, the doping dose and the cap thickness. The cap thickness was varied from 1 nm to 5 nm, and the doping dose was varied from $1\times10^{14}$ cm$^{-3}$ to $1\times10^{17}$ cm$^{-3}$. The transistor gate length is maintained at 0.18 $\mu$m and the source-drain separation at 1.5 $\mu$m. In general, it is found from Fig. 6.36 that the maximum transconductance increases as cap thickness decreases and it lies between 250 to 300 mS/mm.

Fig. 6.37 demonstrates the effect of increasing the cap doping level for devices having a fixed cap thickness of 5 nm. The peak cutoff frequency is found to insensitive to cap doping beyond a dose of about $1\times10^{14}$ cm$^{-3}$. The

gain of the devices remains almost constant, over the range of parameters considered, at around 20. For thicker cap layers, the peak is reached at a lower doping. The maximum value of $g_m$ is found to be around 400 mS/mm.

The impact of buffer layer thickness on the cutoff frequency of Schottky gate devices is shown in Fig. 6.38. Simulation results indicate that it could not be raised above 47 GHz for gate length 0.18 $\mu$m. In Fig. 6.39, cutoff frequency as a function of drain voltage is plotted for $\delta$-doping concentration in the range $2\times10^{18}cm^{-3}$ to $6\times10^{18}cm^{-3}$, while in Fig. 6.40 the same for $\delta$-layer thickness are given. Fig. 6.39 shows a decrease in the cutoff frequency as the frontside and backside $\delta$-doping level is reduced. When the thickness of frontside $\delta$-doped layer is decreased to 2 nm, the peak shifts to a higher gate bias and also the maximum value decreases, while the reduction of backside $\delta$-doped layer has a little effect. In Fig. 6.41, the transconductance as a function of gate voltage is plotted for both backside and frontside spacer layer thickness in the range 1 nm to 2 nm.

For microwave applications, low noise is important. An empirical figure of merit for noise was first introduced by Fukui [56]. It suggests that noise decreases as $\frac{\sqrt{g_m}}{f_t}$. This will favor the HEMT devices, which obtain their high cutoff frequency with a low transconductance. Other important parameters are the source and gate resistances. These are likely to be the same for all of the devices considered. To achieve low gate capacitance, a metal T-gate must be used. For the Schottky gate HFETs, typical values are 0.1-0.2 fF/$\mu$m. The SiGe technology is compatible with conventional CMOS processing [5, 57], which could be used for much of the fabrication of these devices, although the strained-Si channel may be degraded, if the processing thermal budget is too high.

## 6.7 Summary

As challenges to downscaling continue to grow, it is important to examine nonconventional avenues for improving CMOS performance. Because hole mobility is lower than electron mobility in Si, it would be attractive to use materials that could enhance hole over electron mobility, and thus, provide a balanced CMOS device that reduces the total area used for a circuit, as well as improves its performance. In this chapter, we have presented results and discussed on buried-channel $Si_{1-x}Ge_x$ and $Si_{1-x-y}Ge_xC_y$ p-MOSFETs. SiGe-channel p-MOSFETs on SIMOX substrates have been analyzed and compared to SiGe-bulk devices. It is observed that the SiGe-SIMOX devices exhibit improved saturation current drive and channel saturation transconductance over that of SiGe-bulk devices. These predictive simulation results illustrate the availability of a new degree of freedom in SiGe p-HFET design on SOI

substrates.

A physics-based, semi-analytical, charge control model for $Si_{1-x}Ge_x$ quantum well MOS structures was used for the purposes of efficient device simulation. This model adequately describes the onset of carrier saturation in the well and various inversion regimes in the Si cap without the use of ad-hoc fitting parameters. The role of a $p^+$ delta doping layer was found to be useful for threshold adjustment and improvement of the transconductance. Finally, the high Ge content devices are susceptible to large subthreshold leakage. This is a result of hole spillover from the source/drain regions into the lower energy region of the $Si_{1-x}Ge_x$ quantum well. This spillover can be suppressed through the introduction of a $p^+$-doping spike below the $Si_{1-x}Ge_x$ quantum well. For ULSI circuit requirements, the $p^+$-doping spike becomes a necessary feature of the p-MOSFETs. It has been shown that with even modest amounts of Ge (10% to 20%), buried-channel $Si_{1-x}Ge_x$ p-MOSFETs can be used in deeply scaled devices and provide performance enhancement over Si. Partially strain-compensated $Si_{1-x-y}Ge_xC_y$ may be a viable option for thermal stability. Design guidelines for SiGe p-MOSFETs have been presented. A 2D Virtual Wafer Fabrication (VWF) simulation suite has been employed for the technology CAD (TCAD) of SiGe channel p-MOSFETs.

Computer simulation has been used to compare the high frequency performance of strained-Si/SiGe/Si-based heterojunction HEMTs having different geometries. The rf performance for the self-aligned HEMTs is highest, varying from 10 GHz for the 1.8 $\mu$m device to 100 GHz for the 0.1 $\mu$m device. This comparison suggests that the best overall performance can be achieved with the self-aligned HEMT. Even though the highest transconductances are obtained with the Schottky gate HFET devices, this is only because of the shorter effective distance between the gate and channel in these devices. Moreover, their higher gate capacitance prevents this advantage from being exploited in high-frequency performance.

## References

[1] D. N. Ngoc, D. L. Harame, J. C. Malinowski, S. J. Jeng, K. T. Schonenberg, M. M. Gilbert, G. D. Berg, S. Wu, M. Soyuer, K. J. Stein, R. A. Groves, S. Subbanna, D. B. Colavito, D. A. Sunderland, and B. S. Meyerson, "A 200 mm SiGe-HBT BiCMOS technology for mixed signal applications," *IEEE BCTM Proc.*, pp. 89–92, 1995.

[2] P. Chevalier, B. Barbalat, L. Rubaldo, B. Vandelle, D. Dutartre, P. Bouillon, T. Jagueneau, C. Richard, F. Saguin, A. Margain, and A. Chantre, "300 GHz $f_{max}$ self-aligned SiGeC HBT optimized towards CMOS compabitility," *IEEE BCTM Dig.*, pp. 120–123, 2005.

[3] D. K. Nayak, J. C. Woo, J. S. Park, K. L. Wang, and K. P. MacWilliams, "Enhancement-mode quantum-well $Ge_xSi_{1-x}$ PMOS," *IEEE Electron Dev. Lett.*, vol. 12, pp. 154–156, 1991.

[4] K. Ismail, "Si/SiGe high-speed field-effect transistors," *IEEE IEDM Tech. Dig.*, pp. 509–512, 1995.

[5] U. Konig, M. Zeuner, G. Hock, T. Hackbarth, M. Gluck, T. Ostermann, and M. Saxarra, "n- and p-type SiGe HFETs and circuits," *Solid-State Electron.*, vol. 43, pp. 1383–1388, 1999.

[6] S. V. Vandebroek, E. Crabbe, B. Meyerson, D. Harame, P. Restle, J. Stork, and J. Johnson, "SiGe-channel heterojunction p-MOSFETs," *IEEE Trans. Electron Dev.*, vol. 41, pp. 90–101, 1994.

[7] D. K. Nayak, J. C. S. Woo, G. K. Yabiku, K. P. MacWilliams, J. S. Park, and K. L. Wang, "High-mobility GeSi PMOS on SIMOX," *IEEE Electron Dev. Lett.*, vol. 14, pp. 520–522, 1993.

[8] M. Yoshimi, M. Terauchi, A. Nishiyama, O. Arisumi, A. Murakoshi, K. Matsuzawa, N. Shigyo, S. Takeno, M. Tomita, K. Suzuki, Y. Ushiku, and H. Tango, "Suppression of the floating-body effect in SOI MOSFETs by the bandgap engineering method using a $Si_{1-x}Ge_x$ source structure," *IEEE Trans. Electron Dev.*, vol. 44, pp. 423–429, 1997.

[9] S. J. Mathew, G. Niu, W. B. Dubbelday, and J. D. Cressler, "Characterization and profile optimization of SiGe pFETs on silicon-on-sapphire," *IEEE Trans. Electron Dev.*, vol. 46, pp. 2323–2332, 1999.

[10] Y. C. Yeo, V. Subramanian, J. Kedzierski, P. Xuan, T.-J. King, J. Bokor, and C. Hu, "Nanoscale ultra-thin-body silicon-on-insulator P-MOSFET with a SiGe/Si heterostructure channel," *IEEE Electron Dev. Lett.*, vol. 21, pp. 161–163, 2000.

[11] R. People and J. C. Bean, "Band alignments of coherently strained $Ge_xSi_{1-x}$/Si heterostructures on <001> $Ge_ySi_{1-y}$ substrates," *Appl. Phys. Lett.*, vol. 48, pp. 538–540, 1986.

[12] J. Welser, J. L. Hoyt, and J. F. Gibbons, "Electron mobility enhancement in strained-Si N-type metal-oxide-semiconductor field-effect transistors," *IEEE Electron Dev. Lett.*, vol. 15, pp. 100–102, 1994.

[13] S. E. Swirhun, Y. H. Kwark, and R. M. Swanson, "Measurement of electron lifetime, electron mobility and bandgap narrowing in heavily doped p-type silicon," *IEEE IEDM Tech. Dig.*, pp. 24–27, 1986.

[14] D. B. M. Klaassen, J. M. Slotboom, and H. C. De Graaff, "Unified apparent bandgap narrowing in n- and p-type silicon," *Solid-State Electron.*, vol. 35, pp. 125–129, 1992.

[15] J. W. Slotboom and H. C. deGraff, "Measurement of bandgap narrowing in Si bipolar transistor," *Solid-State Electron.*, vol. 19, pp. 857–862, 1976.

[16] Th. Vogelsang and K. R. Hofmann, "Electron transport in strained-Si layers on $Si_{1-x}Ge_x$ substrates," *Appl. Phys. Lett.*, vol. 63, pp. 186–188, 1993.

[17] T. Manku and A. Nathan, "Electron drift mobility model for devices based on unstrained and coherently strained-$Si_{1-x}Ge_x$ grown on <001> silicon substrate," *IEEE Trans. Electron Dev.*, vol. 39, pp. 2082–2089, 1992.

[18] D. M. Caughey and R. E. Thomas, "Carrier mobilities in silicon empirically related to doping and field," *IEEE Proceedings*, vol. 55, pp. 2192–2193, 1967.

[19] H. Miyata, T. Yamada, and D. K. Ferry, "Electron transport properties of a strained-Si layer on a relaxed $Si_{1-x}Ge_x$ substrate by Monte Carlo simulation," *Appl. Phys. Lett.*, vol. 62, pp. 2661–2663, 1993.

[20] T. Yamada, J.-R. Zhou, H. Miyata, and D. K. Ferry, "Velocity overshoot in a modulation-doped $Si/Si_{1-x}Ge_x$ structure," *Semicond. Sci. Technol.*, vol. 9, pp. 775–777, 1994.

[21] D. J. Harame, J. H. Comfort, J. D. Cressler, E. F. Crabbe, J. Y. C. Sun, B. S. Meyerson, and T. Tice, "Si/SiGe epitaxial-base transistors: Part I: materials, physics, and circuits," *IEEE Trans. Electron Dev.*, vol. 42, pp. 455–467, 1995.

[22] K. K. Liou, P. Mei, U. Gennser, and E. S. Yang, "Effects of Ge concentration on SiGe oxidation behavior," *Appl. Phys. Lett.*, vol. 59, pp. 1200–1202, 1991.

[23] J. Eugene, F. K. LeGoues, V. P. Kesan, S. S. Iyer, and F. M. d'Heurle, "Diffusion versus oxidation rates in silicon-germanium alloys," *Appl. Phys. Lett.*, vol. 59, pp. 78–80, 1991.

[24] S. S. Iyer, P. M. Solomon, V. P. Kesan, A. A. Bright, J. L. Freeouf, T. N. Nguyen, and A. C. Warren, "A gate-quality dielectric system for SiGe metal-oxide-semiconductor devices," *IEEE Electron Dev. Lett.*, vol. EDL-12, pp. 246–248, 1991.

[25] M. Mukhopadhyay, S. K. Ray, D. K. Nayak, and C. K. Maiti, "Ultrathin oxides using $N_2O$ on strained-$Si_{1-x}Ge_x$," *Appl. Phys. Lett.*, vol. 68, pp. 1262–1264, 1996.

[26] P. W. Li, H. K. Liou, E. S. Yang, S. S. Iyer, T. P. Smith, and Z. Lu, "Formation of stoichiometric SiGe oxide by electron cyclotron resonance plasma," *Appl. Phys. Lett.*, vol. 60, pp. 3265–3267, 1992.

[27] P. W. Li, E. S. Yang, Y. F. Yang, J. O. Chu, and B. S. Meyerson, "SiGe pMOSFETs with gate oxide fabricated by microwave electron cyclotron resonance plasma processing," *IEEE Electron Dev. Lett.*, vol. 15, pp. 402–405, 1994.

[28] A. Sadek, K. Ismail, M. A. Armstrong, D. A. Antoniadis, and F. Stern, "Design of Si/SiGe heterojunction complementary metal-oxide-semiconductor transistors," *IEEE Trans. Electron Dev.*, vol. 43, pp. 1224–1232, 1996.

[29] V. P. Kesan, S. Subbanna, P. J. Restle, M. J. Tejwani, J. M. Aitken, S. S. Iyer, and J. A. Ott, "High performance 0.25 $\mu$m p-MOSFETs with silicon germanium channels for 300 K and 77 K operation," *IEEE IEDM Tech. Dig.*, pp. 25–28, 1991.

[30] P. Bouillon, T. Skotnicki, C. Kelaidis, R. Gwoziecki, P. Dollfus, J.-L. Regolini, I. Sagnes, and S. Bodnar, "Search for the optimal channel architecture for 0.18/0.12 $\mu$m bulk CMOS experimental study," *IEEE IEDM Tech. Dig.*, pp. 559–562, 1996.

[31] E. Murakami, K. Nakagawa, A. Nishida, and M. Miyao, "Fabrication of strain-controlled SiGe/Ge MODFET with ultrahigh hole mobility," *IEEE Trans. Electron Dev.*, vol. 41, pp. 857–861, 1994.

[32] R. People, "Physics and applications of $Ge_xSi_{1-x}$ strained-layer heterostructures," *IEEE J. Quantum Elec.*, vol. 22, pp. 1696–1710, 1986.

[33] D. K. Nayak, J. C. S. Woo, J. S. Park, K. L. Wang, and K. P. MacWilliams, "High-mobility p-channel metal-oxide-semiconductor field-effect transistor on strained-Si," *Appl. Phys. Lett.*, vol. 62, pp. 2853–2855, 1993.

[34] C. K. Maiti, L. K. Bera, and S. Chattopadhyay, "Strained-Si heterostructure field effect transistors," *Semicond. Sci. Technol.*, vol. 13, pp. 1225–1246, 1998.

[35] K. Bhaumik and Y. S. Diamand, "Semi-analytical model for charge control in SiGe quantum well MOS structures," *Solid-State Electron.*, vol. 36, pp. 961–968, 1993.

[36] B. K. Ip and J. R. Brews, "Design tradeoff guidelines for n-channel delta-doped Si/SiGe heterojunction MOSFETs," *Solid-State Electron.*, vol. 44, pp. 593–604, 2000.

[37] S. K. Chun and K. L. Wang, "Effective mass and mobility of holes in strained-$Si_{1-x}Ge_x$ layers on (001) $Si_{1-y}Ge_y$ substrate," *IEEE Trans. Electron Dev.*, vol. 39, pp. 2153–2164, 1992.

[38] T. Manku and A. Nathan, "Effective mass for strained p-type $Si_{1-x}Ge_x$," *J. Appl. Phys.*, vol. 69, pp. 8414–8416, 1991.

[39] F. Stern, "Self-consistent results for n-type Si inversion layers," *Phys. Rev. B*, vol. 5, pp. 4891–4899, 1972.

[40] S. Maikap, B. Senapati, and C. K. Maiti, "Technology CAD of SiGe-heterojunction field effect transistors," *Def. Sci. J.*, vol. 51, pp. 195–199, 2001.

[41] S. Maikap, L. K. Bera, S. K. Ray, S. John, S. K. Banerjee, and C. K. Maiti, "Electrical characterization of Si/$Si_{1-x}Ge_x$/Si quantum well heterostructures using a MOS capacitor," *Solid-State Electron.*, vol. 44, pp. 1029–1034, 2000.

[42] S. K. Ray, S. John, S. Oswal, and S. K. Banerjee, "Novel SiGeC channel heterojunction PMOSFET," *IEEE IEDM Tech. Dig.*, pp. 261–264, 1996.

[43] A. D. Sadovnikov, A. Nathan, and T. Manku, "Numerical simulation of transconductance in Si/SiGe p-MOSFETs," *Solid-State Electron.*, vol. 39, pp. 1199–1203, 1996.

[44] P. M. Garone, V. Venkataraman, and J. C. Sturm, "Mobility enhancement and quantum mechanical modeling in $Ge_xSi_{1-x}$ channel MOSFETs from 90 to 300 K," *IEEE IEDM Tech. Dig.*, pp. 29–32, 1991.

[45] Z. Shi, X. Chen, D. Onsongo, E. J. Quinones, and S. K. Banerjee, "Simulation and optimization of strained-$Si_{1-x}Ge_x$ buried channel p-MOSFETs," *Solid-State Electron.*, vol. 44, pp. 1223–1228, 2000.

[46] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors." San Jose, Calif., 2005.

[47] C. K. Maiti and G. A. Armstrong, *Applications of Silicon-Germanium Heterostructure Devices.* Institute of Physics Publishing, UK, 2001.

[48] S. C. Jain, H. J. Osten, B. Dietrich, and H. Rucker, "Growth and properties of strained-$Si_{1-x-y}Ge_xC_y$ layers," *Semicond. Sci. Technol.*, vol. 10, pp. 1289–1302, 1995.

[49] W. Liu, A. St. Amour, J. C. Sturm, Y. R. J. Lacroix, M. L. W. Thewalt, C. W. Magee, and D. Eaglesham, "Growth and photoluminescence of high quality SiGeC random alloys on silicon substrates," *J. Appl. Phys.*, vol. 80, pp. 3043–3047, 1996.

[50] E. J. Quinones, S. John, S. K. Ray, and S. K. Banerjee, "Design, fabrication, and analysis of SiGeC heterojunction PMOSFETs," *IEEE Trans. Electron Dev.*, vol. 47, pp. 1715–1725, 2000.

[51] T. Ngai, W. J. Qi, R. Sharma, J. L. Fretwell, X. Chen, J. C. Lee, and S. K. Banerjee, "Transconductance improvement in surface-channel SiGe p-metal-oxide-silicon field-effect transistors using a $ZrO_2$ gate dielectric," *Appl. Phys. Lett.*, vol. 78, pp. 3085–3087, 2001.

[52] Z. Shi, D. Onsongo, K. Onishi, J. C. Lee, and S. K. Banerjee, "Mobility enhancement in surface channel SiGe PMOSFETs with $HfO_2$ gate dielectrics," *IEEE Electron Dev. Lett.*, vol. 24, pp. 34–36, 2003.

[53] B. Senapati, S. K. Samanta, and C. K. Maiti, "Simulation of strained-Si/SiGe HEMTs," *Proc. IWPSD-2003*, pp. 833–835, 2003.

[54] K. Ismail, J. O. Chu, K. Chan, and B. S. Mayerson, "High performance Si/SiGe n-type modulation doped transistors," *IEEE Electron Dev. Lett.*, vol. 14, pp. 348–350, 1993.

[55] M. Gluck, T. Hackbarth, U. Konig, A. Haas, G. Hock, and E. Kohn, "High $f_{max}$ n-type Si/SiGe MODFETs," *Electron. Lett.*, vol. 33, pp. 335–337, 1997.

[56] H. Fukui, "Optimum noise figure of microwave GaAs MESFETs," *IEEE Trans. Electron Dev.*, vol. 26, pp. 1032–1037, 1979.

[57] A. Sadek and K. Ismail, "Si/SiGe CMOS possibilities," *Solid-State Electron.*, vol. 38, pp. 1731–1734, 1995.

# 7

# *Strained-Si Heterostructure MOSFETs*

As the end of the Si roadmap is approached, strained-Si channels offer a way to compensate for the reduction in MOSFET drive currents that accompany transistor scaling. Strain is an effective mechanism to modify the carrier transport properties of silicon. Strained-Si/SiGe MOSFETs take advantage of strain-induced changes of carrier transport in silicon and obtain current drive enhancements. The field has centered on process-induced strained-Si MOSFETs, because of its better compatibility with mainstream CMOS, and the impressive mobility enhancements that can be realized in that system at aggressively scaled gate lengths. Jankovic and O'Neill [1] have investigated the influence of strained-Si cap layers on n-p-n heterojunction bipolar transistors (HBTs) fabricated on virtual substrates. Using an approximate theoretical model, it has been shown that the presence of a strained-Si/SiGe (relaxed) heterojunction barrier in the emitter can substantially improve the current gain, relaxing the need for a high Ge content in the strained-SiGe base. Furthermore, two-dimensional numerical simulations of a virtual substrate HBT with a realistic geometry demonstrate that, besides the current gain enhancement, a three-times improvement in $f_t$ and $f_{max}$ can be realized when a strained-Si/SiGe emitter is incorporated. Although high performance complementary heterojunction MOSFETs (HCMOS) using strained-Si channels fabricated on virtual substrates have been demonstrated [2], it is important that heterojunction bipolar transistors (HBTs) are also integrated together with HFETs for future Bi-HCMOS technology on virtual SiGe substrates.

In this chapter, MOSFET performance enhancement by using strain will be discussed. Various types of stress used in leading-edge advanced CMOS technology will be discussed in depth for reported experimental devices. The emphasis will be to highlight the issues which the industry needs to account for in developing new process and devices. We quantitatively explain the experimental results and show that the larger and undesirable threshold voltage shift for biaxial stress is fundamental to biaxial strain. An overview of strained-Si process technology and process integration issues in Si CMOS technology is presented.

## 7.1 Operating Principle

Since in earlier chapters we have provided overviews of the transport, materials and bandstructure issues of strained-Si and SiGe, in this chapter we will focus on their application in heterostructure FETs. We will briefly discuss the transport issues of strained-Si where necessary that are relevant to HFETs. Then we will discuss their applications in buried- and surface-channel HFETs, with an emphasis on both n- and p-HFETs. Observation of high electron mobility in two-dimensional electron gas formed in strained-Si was reported in early 1990s [3]. Application of strained-Si in MOSFET channel was proposed by Keyes as early as in 1986 [4], and the first demonstration of strained-Si MOSFET was reported by Welser et al. [5] using relaxed graded-SiGe buffer layer.

For strained-Si/SiGe channel MOSFETs, the critical device parameters can be identified by first examining the drain current expression in a long-channel MOSFET which need to be modified due to strain. The classical MOSFET equations for a long-channel device operating in linear and saturation regions are given by:

$$I_{d,lin} = \frac{\mu C_{ox} W}{L}\left[(V_{gs} - V_t)V_{ds} - V_{ds}^2/2\right] \tag{7.1}$$

$$I_{d,sat} = \frac{\mu C_{ox} W}{2L}\left(V_{gs} - V_t\right)^2 \tag{7.2}$$

where W and L are the gate width and length, $\mu_{eff}$ is the effective channel mobility, $C_{ox}$, the gate dielectric capacitance per unit area, $V_t$ is the threshold voltage, and $V_{gs}$ and $V_d$ are the gate and drain biases, respectively.

In the nanometer realm, the carrier transport in the device becomes ballistic and Eqns. 7.1 and 7.2 becomes dependent on the carrier velocity as:

$$I_{d,sat} = C_{ox} W < v(0) > \left(V_{gs} - V_t\right) \tag{7.3}$$

Hence, it can be seen that the square-law relationship approaches linearity in a ballistic model. In the above equation, $< v(0) >$ is the average carrier velocity at source, which is related to the mobility by Eqns. 7.1 and 7.2. As the channel length approaches zero, $<v(0)>$ is the unidirectional thermal velocity $v_t$, which is given by [6]:

$$v_t = \sqrt{\frac{2k_B T_L}{\pi m^*}} \tag{7.4}$$

The drain current in the nanoscale limit can now be equated as:

$$I_{d,sat} = C_{ox} W \sqrt{\frac{2k_B T_L}{\pi m^*}} \left( V_{gs} - V_t \right) \tag{7.5}$$

Therefore, it is obvious that improving the mobility by application of strain enhances the performance of the device. In order to realize the strain-enhanced CMOS performance to the fullest extent in ULSI applications, both fundamental and technological challenges need to be addressed. Strained-Si technology is being adopted for two key reasons. First, the processing, engineering, and design costs associated with implementing strained-Si are modest. Second, strained-Si offers performance enhancement with 30-60% product level improvement in present-day production on 300-mm wafers. The additional processing cost is modest at $\sim$ 1-3%.

### 7.1.1 Threshold Voltage

The threshold voltage of MOSFET, which has always been the major concern of device scaling, is closely related to device structure parameters, e.g., oxide thickness, junction depth, gate length, and terminal voltages. The increasing complexity of modern MOS device structure requires clear understanding on the behavior of threshold voltage. Understanding the magnitude of the threshold voltage shift is important when determining the performance gain of strained-Si. Therefore, many efforts have been devoted to the investigation of this electrical parameter. Since performance benchmarking needs to be done at constant off-state leakage, an adjustment to compensate for the strain-induced threshold voltage shift is required. The threshold voltage expression for biaxial tensile-stressed Si on relaxed-$Si_{1-x}Ge_x$ MOSFETs has been derived [7]. The same derivation can be used for uniaxial process stress, except strain-induced changes in the $n^+$-polySi gate electron affinity and bandgap need to be included. Large differences in the experimentally observed strain-induced threshold voltage shifts for uniaxial and biaxial tensile-stressed silicon n-channel MOSFETs have been explained and quantified [8]. Using the deformation potential theory, key quantities that affect threshold voltage (electron affinity, bandgap, and valence band density of states) are expressed as a function of strain. Unlike wafer substrate strain techniques, an interesting feature of process strain, such as a tensile nitride capping-layer or SiGe in the source and drain [9], is that the Si channel and gate are both strained as shown in Fig. 7.1.

To quantify the strain-induced n-MOSFET threshold voltage shift, deformation potential theory is used to calculate electron affinity, bandgap narrowing, and density of states as a function of biaxial and uniaxial stress. The threshold voltage shift for biaxial strain is given by [10]

$$q\Delta V_{th}(\sigma) = \Delta E_c(\sigma) + (m-1)\left[\Delta E_c(\sigma) + kTln\frac{N_v(0)}{N_v(\sigma)}\right] \tag{7.6}$$

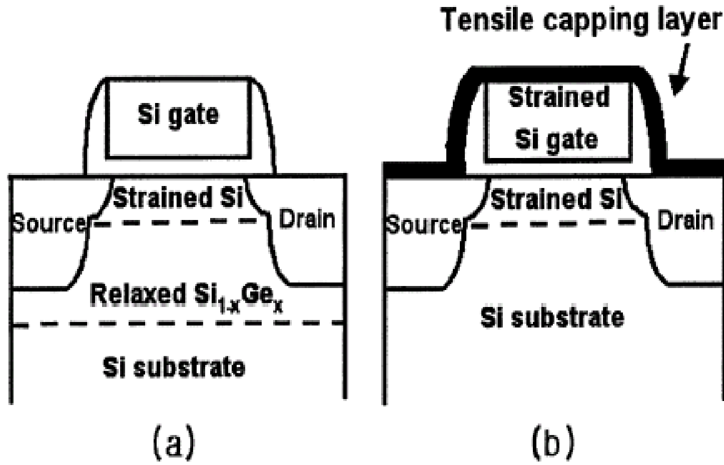



**FIGURE 7.1**
(a) Biaxial and (b) uniaxial tensile-stressed MOSFET. After J.-S. Lim et al., *IEEE Electron Dev. Lett.*, Vol. 25, 2004(731-733). With permission.

**TABLE 7.1**
Si electron affinity and bandgap narrowing calculated as a function of strain. After S. E. Thompson et al., *IEEE IEDM Tech. Dig.*, 2004(221-224). With permission.

| Stress type | Deformation potential constants (eV) | Deformation potential constants (eV) | $\Delta E_c(\sigma)$ (eV) | $\Delta E_g(\sigma)$ (eV) |
|---|---|---|---|---|
| [110] Uniaxial | $\Xi_d = 1.13$<br>$\Xi_u = 9.16$ | a = 2.46<br>b = -2.35<br>d = -5.08 | $-2.66\epsilon$ | $-6.19\epsilon$ |
| Biaxial (1) | $\Xi_d = 1.13$<br>$\Xi_u = 9.16$ | a = 2.46<br>b = -2.35 | $-5.67\epsilon$ | $-17.01\epsilon$ |
| Biaxial (2) | $\Xi_d = -10.7$<br>$\Xi_u = 10.5$ | a = -9.7<br>b = -2.33 | $-41.04\epsilon$ | $-44.72\epsilon$ |

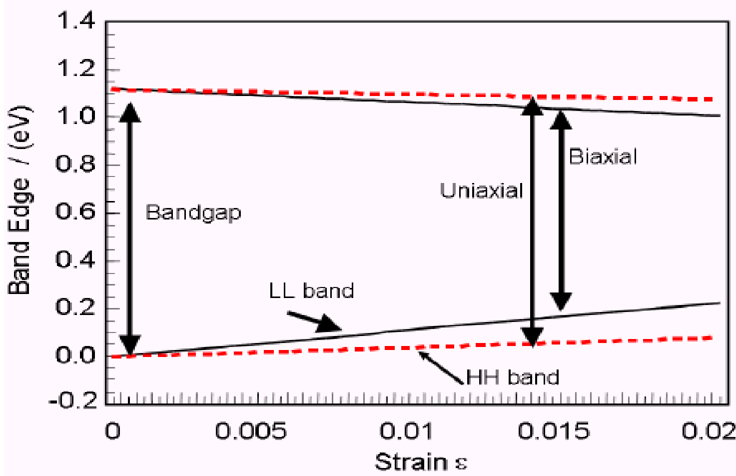



**FIGURE 7.2**
Si conduction and valence band shift for uniaxial and biaxial stress computed using deformation potential theory. After S. E. Thompson et al., *IEEE IEDM Tech. Dig.*, 2004(221-224). With permission.

and for uniaxial stress it is given by

$$q\Delta V_{th}(\sigma) = (m-1)\left[\Delta E_c(\sigma) + kTln\frac{N_v(0)}{N_v(\sigma)}\right] \tag{7.7}$$

where m is the body coefficient, $E_c$ the conduction band edge, $E_g$ the energy gap and $N_v$ the valence band density of states.

The larger threshold shift for biaxial strain results from two effects. First, biaxial tensile strain causes larger Si channel bandgap narrowing as shown in Fig. 7.2. The larger bandgap narrowing results for several reasons, one being a larger shift in the valence band edge for biaxial than uniaxial stress. This occurs since the light (vs. heavy) hole band shifts more and the valence band edge is set by the light hole band for biaxial tensile stress while being set by the heavy hole for uniaxial tensile stress (not to be confused with uniaxial compressive stress where the light hole band is the top band).

Second, uniaxial process-induced strain does not have a $\Delta E_c$ term since the gate is also strained. The $V_{th}$ shift is calculated using the above equations for two sets of deformation potentials and shown in Fig. 7.3. The two sets of deformation potentials are used to bracket the large uncertainty in this material property. The calculated threshold voltage shift is at least four times larger for biaxial than uniaxial stress. The $V_{th}$ shift is in agreement with uniaxial wafer bending as published biaxial strained-Si on relaxed-SiGe MOSFET experimental data [11, 12, 13] included in Fig. 7.3.

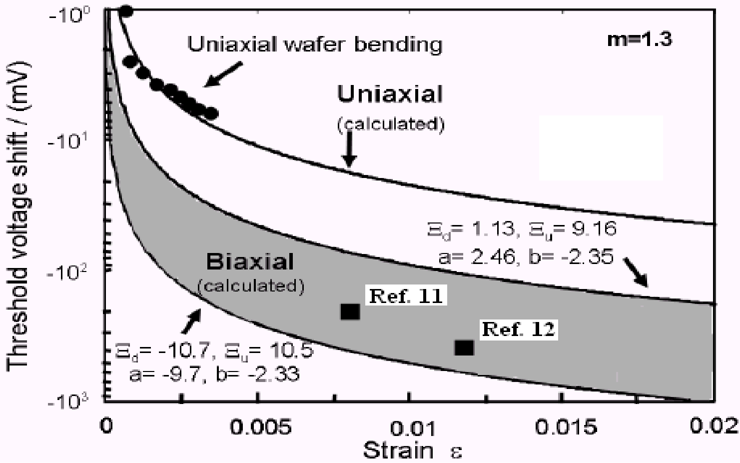



**FIGURE 7.3**
Computed and measured threshold voltage shift for n-MOSFETs under uniaxial and biaxial stress. Shift is about 4× for biaxial strain. After S. E. Thompson et al., *IEEE IEDM Tech. Dig.*, 2004(221-224). With permission.

The calculated threshold voltage shift has been found to be in agreement with uniaxial wafer bending and published biaxial strained-Si on relaxed-$Si_{1-x}Ge_x$ experimental data [14, 15] and explains the technologically important observation of a significantly larger (4×) threshold voltage shift for biaxial relative to uniaxial stressed MOSFETs. The large threshold shift for biaxial stress is shown to result from the stress-induced change in the Si channel electron affinity and bandgap. The small threshold voltage shift for uniaxial process tensile stress is shown to result from the $n^+$-polySi gate in addition to the Si channel being strained and significantly less bandgap narrowing.

As discussed in chapter 4, strained-Si offers both valence and conduction band offsets. In strained-Si n-MOSFETs, the negative valence band offset makes the Fermi level closer to the conduction band; in turn, it induces more electrons in the inversion layer for the same gate bias. Thus, the band offset lowers the threshold and makes the channel depletion shallower. As the valence band offset is linearly proportional to the Ge content, each 10% increase of Ge results in about a -100 mV threshold shift and 4% degradation of the subthreshold slope in n-MOSFETs.

Biaxially strained-Si MOSFET technology, yielding strained-Si channels on SiGe buffer layers, has attracted much interest because of the potential device performance enhancement due to improved carrier transport properties. However, lower bandgap tends to have a reduced threshold voltage and its adjustment via redesign of the device undermines the actual performance enhancement that can be achieved. Heavier channel doping, for example, can

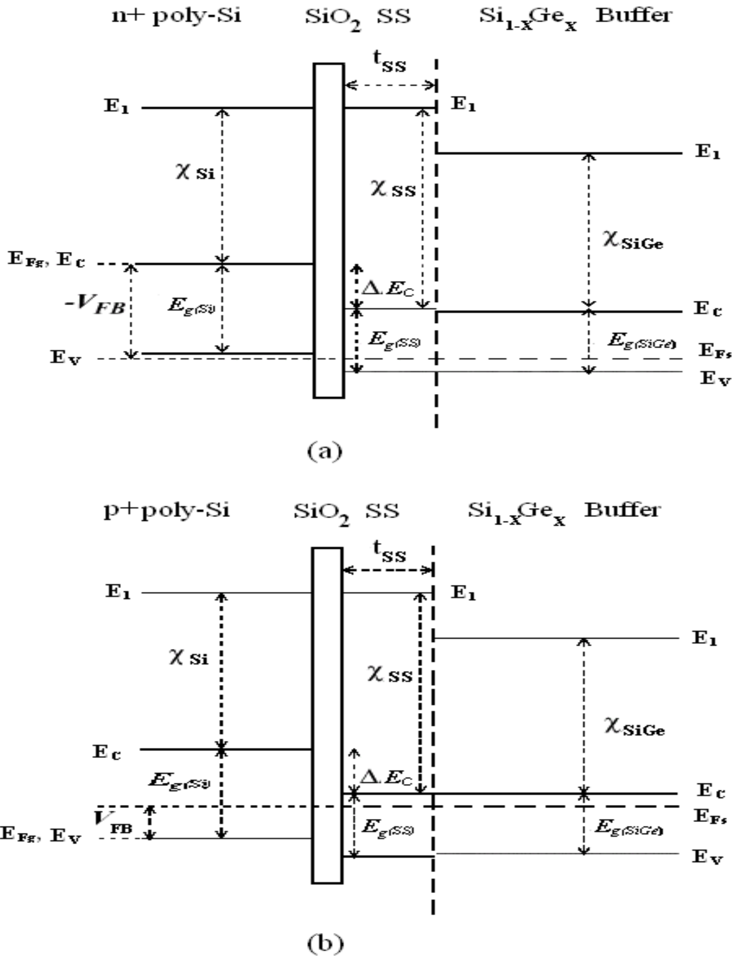



**FIGURE 7.4**
Energy band diagrams at the flat-band condition for strained-Si/$Si_{1-x}Ge_x$: (a) n-MOSFETs and (b) p-MOSFETs, where $\Delta E_c = \chi_{ss}$-$\chi_{Si} = 0.57$x (eV) and $E_{g,ss} = E_{g,Si}$ - 0.4x (eV); $E_i$ represents the local vacuum or reference energy level, and $E_{Fg}$ and $E_{Fs}$ are the Fermi levels in the gate and Si substrate, respectively. The differences between $N_c$ and $N_v$ of the strained-Si layer and the SiGe buffer layer are ignored. After W. Zhang and J. G. Fossum, *IEEE Trans. Electron Dev.*, Vol. 52, 2005(263-268). With permission.

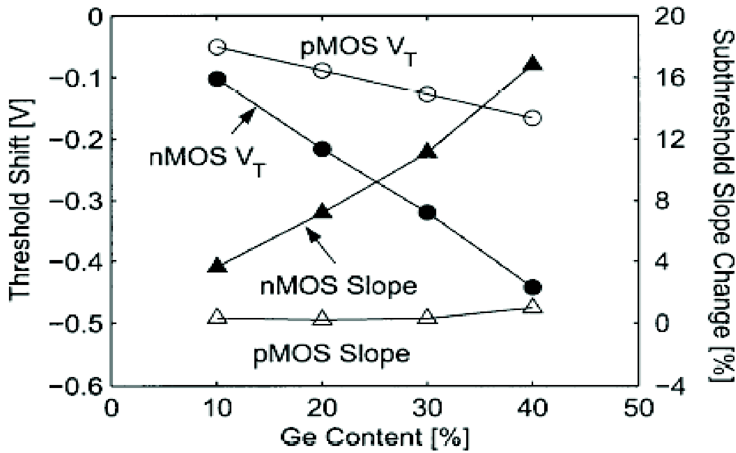



**FIGURE 7.5**
Simulated threshold voltage shifts (circles) and subthreshold slope changes (triangles), as a function of Ge content in the relaxed-SiGe layer. The thickness of the strained-Si layer is 10 nm, the gate length is 1.0 $\mu$m, and the drain bias is 50 mV. After J.-S. Goo et al., *IEEE Electron Dev. Lett.*, Vol. 24, 2003(568-570). With permission.

be easily done to compensate for reduced $V_t$, but such design degrades the enhanced carrier mobility and increases source/drain junction capacitance, significantly limiting the speed performance improvement. Optimizing the channel doping and/or using metal gate have been utilized in attempts to overcome this problem, but the technological issues are not trivial. To facilitate such design optimization, as well as to reasonably predict the speed performance improvement that can be achieved with strained-Si MOSFET technology, one needs more accurate physical modeling of the $V_t$ shift in terms of the variations of bandgap, electron affinity, and even effective conduction and valence band densities of states.

Threshold voltage shifts, $\Delta V_t$ in biaxially strained-Si/SiGe CMOS devices, defined by the shifted 2D energy subbands and modified effective conduction and valance band densities of states, have been studied in detail by Zhang and Fossum [7]. Increased electron affinity as well as bandgap narrowing in the strained-Si layer are shown to be the predominant components of $\Delta V_t$, whereas the density-of-state terms tend to be relatively small but not insignificant. Energy band diagrams at the flat-band condition for strained-Si/$Si_{1-x}Ge_x$ n- and p-MOSFETs are shown in Fig. 7.4.

The correlation between the threshold voltage shift and strained layer thickness has been investigated via device simulation by Goo et al. [13]. Fig. 7.5 presents simulated threshold voltage shifts and subthreshold slope changes of

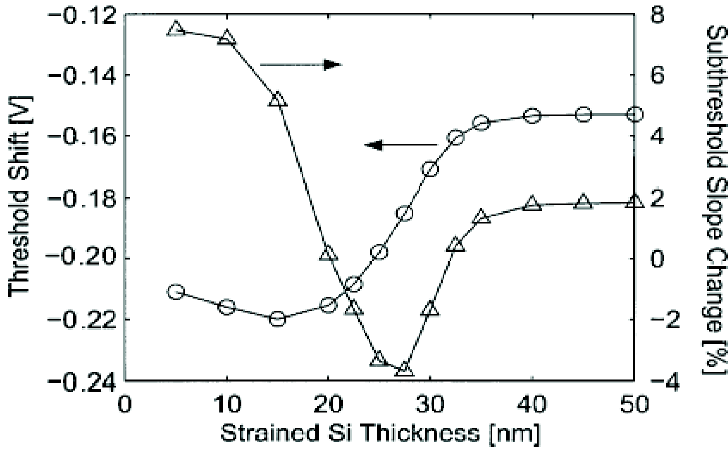



**FIGURE 7.6**
Simulated threshold voltage shifts (circles) and subthreshold slope changes (triangles) of n-MOSFETs, as a function of the thickness of the strained-Si layer. The Ge content in the relaxed-SiGe layer is 20%, the gate length is 1.0 $\mu$m, and the drain bias is 50 mV. After J.-S. Goo et al., *IEEE Electron Dev. Lett.*, Vol. 24, 2003(568-570). With permission.

strained-Si MOSFETs for different amounts of Ge in the relaxed-SiGe layer. It is seen that in a certain range of the strained-Si layer thickness, the threshold and subthreshold slope change gradually, posing a concern of larger device parameter variations. This happens when the depletion edge is close to the boundary of the strained-Si/SiGe heterojunction structure. A larger threshold distribution has been observed in devices fabricated with a strained layer thickness comparable to the depletion depth. This result suggests that more strain makes devices harder to optimize.

The dependence on the thickness of the strained-Si layer is shown in Fig. 7.6. A gradual transition is observed in the range of 15-35 nm strained-Si thickness, both for the threshold shift and subthreshold slope change. It is expected that the device parameter fluctuations are larger in that range of thickness, subject to the strained layer thickness variation. This phenomenon is related to the strained layer thickness and channel depletion depth.

Strained-Si devices with two different strained layer thicknesses were fabricated using an identical standard CMOS process. The strain of Si is obtained from a Si layer grown epitaxially on a relaxed-$Si_{0.80}Ge_{0.20}$ layer. A CMP step was introduced in the middle of SiGe epitaxy to reduce the surface roughness [16]. Overall thermal budget and etch steps were controlled tightly to minimize strain relaxation, Ge outdiffusion, and strained-Si consumption. Due to the germanosilicide problem, NiSi was employed instead of CoSi [17].

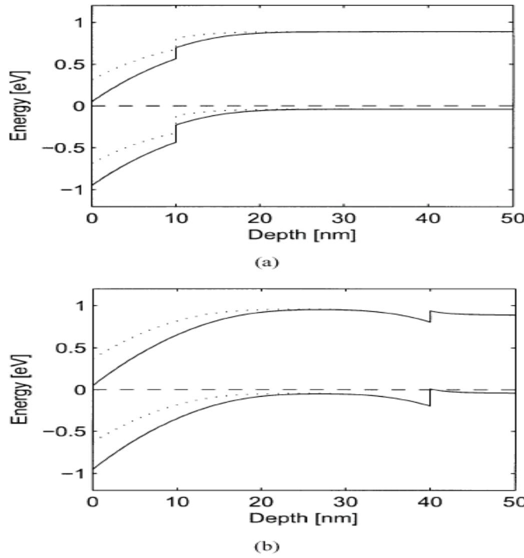



**FIGURE 7.7**
Calculated energy band diagrams for the heterojunction of the strained-Si and 20% SiGe layers for different strained layer thickness. The solid and dotted lines represent $V_{gs} = V_t$ and $V_{gs} = 0$ V, respectively. The dashed line is the Fermi level. (a) 10 nm. (b) 40 nm. After J.-S. Goo et al., *IEEE Electron Dev. Lett.*, Vol. 24, 2003(568-570). With permission.

In contrast, p-MOSFETs exhibit only -40 mV threshold shift per 10% Ge increase, with negligible degradation of the subthreshold slope. This is partly because the conduction band offset is smaller than the valence band offset. Another reason is that the valence band offset acts like a barrier to the holes in p-MOSFETs and forms a parasitic buried channel confining a substantial amount of holes at the interface between the strained-Si and SiGe. As these holes are virtually an extra amount of inversion charges induced by the band structure itself without increasing the gate bias, they make the threshold shift further smaller. A similar phenomenon contributes to the n-MOSFET with a thick strained-Si layer illustrated in Fig. 7.7(b).

Figs. 7.7(a) and (b) compare the band structures for two different strained layer thicknesses. If the thickness is smaller than the depletion depth at zero gate bias, as illustrated in Fig. 7.7(a), the valence band offset fully contributes to the threshold shift. The subthreshold slope degrades due to the higher depletion capacitance, as a consequence of the higher dielectric constant of SiGe and shallower channel depletion depth. On the other extreme, if the strained layer thickness is larger than the maximum depletion depth, the valence band offset partially contributes to the threshold shift and subthreshold slope de-

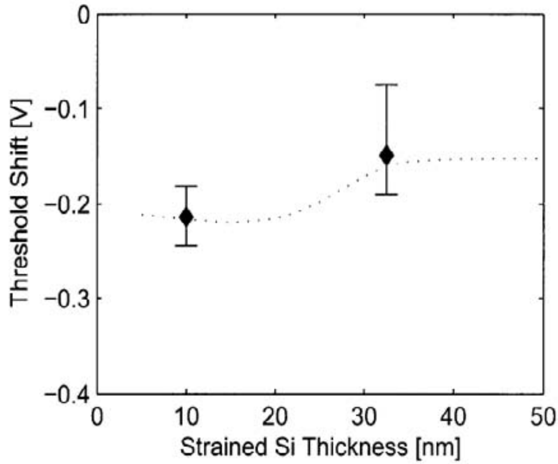



**FIGURE 7.8**
Measured threshold voltage shifts, with respect to control devices, for different strained-Si layer thicknesses. The symbol and bar represent the average and minimum-maximum distribution, respectively. The Ge content in the relaxed-SiGe layer is 20%, the gate length is 1.0 $\mu$m, and the drain bias is 100 mV. The dotted line is MEDICI simulation results presented in Fig. 7.6. After J.-S. Goo et al., *IEEE Electron Dev. Lett.*, Vol. 24, 2003(568-570). With permission.

grades slightly only due to the smaller bandgap of the strained-Si layer.

The parasitic buried channel is also a reason of smaller threshold shift as discussed earlier. In the intermediate range of the thickness, depending on the gate bias, the depletion edge moves across the boundary of the heterojunction structure, thus the threshold becomes different. In such a case, the surface band bending also lowers the band hump in the middle of the strained layer. Like the well-known DIBL effect, this phenomenon results in transfer of extra amount of electrons from the heterojunction valley to the surface channel. Therefore the gate has a better control on the inversion charge. This improves the subthreshold slope as shown in Fig. 7.6.

The depletion depth is subject to scaling, and thus will decrease. On the other hand, the strained layer thickness is not scalable since it needs to be thick enough to confine the wave function - at least 5.7 nm is required to keep the mobility enhancement [16]. This implies that the depletion depth may be comparable to the strained layer thickness at a future technology node, causing worse device uniformity.

Fig. 7.8 shows measured threshold voltage shifts for two different strained layer thicknesses, with respect to the nominal value of conventional Si devices. Both cases showed comparable subthreshold slopes. Some devices with thicker

strained layer showed anomalous characteristics, possibly due to the misfit dislocation, and are excluded. Devices with a thicker strained-Si layer exhibit a smaller threshold shift but with twice as much variation, supporting the earlier discussion.

## 7.2 Uniaxial Stress: Process Flow

A process flow has been developed that allows the strain to be independently controlled for n- and p-MOSFETs by adjusting capping films stress for n-type and Ge content in the source/drain region for p-type. The mismatch in the SiGe to Si lattice causes the p-MOSFETs to be under compressive strain. For small devices the near-complete (if not completely strained due to the lack of misfits) biaxially strained-SiGe in the source/drain region creates significant uniaxial compression in the channel as confirmed by three-dimensional finite element analysis.

Thompson et al. [9] have introduced compressive uniaxial strain in strained-Si p-MOSFETs by inserting a Si recess etch and selective epitaxial $Si_{1-x}Ge_x$ ($x = 0.2$) layer in the source/drain region in a standard CMOS logic technology process flow. The unique advantage of this uniaxial strained-Si process flow is that on the same wafer, compressive strain is introduced into the p- and tensile strain in the n-MOSFETs to improve both the electron and hole mobility. Furthermore, by confining the SiGe to the source/drain and introducing it later in the process flow, the integration challenges become simpler than the biaxial approach and MOSFET self-heating is unchanged [9].

Ghani et al. [18] have reported a combined technique of selective SiGe source-drain and high-stress silicon nitride capping layer, which is a low-cost and highly manufacturable process to induce strain in transistors and allow for separate optimization of pMOS and nMOS devices. The strained pMOS transistor structure features an epitaxially grown compressively strained-SiGe film embedded in the source drain regions. Fig. 7.9 shows a TEM cross-section of the pMOS transistor. A combination of compressive SiGe strain and embedded SiGe S/D geometry induces a large uniaxial compressive strain in the channel region, thereby resulting in significant hole mobility improvement. Dramatic ($> 50\%$) strain-induced hole mobility enhancement is due to channel strain. As shown in Fig. 7.10, the uniaxial strain for this pMOS device results in higher mobility enhancement vs. vertical electric field relative to a biaxially strained device.

Fig. 7.11 shows a TEM cross-section of an nMOS transistor. The unique feature of this nMOS structure is the integration of a post-salicide "highly-tensile" silicon nitride capping layer. The stress from this capping layer is uniaxially transferred to the nMOS channel through the source-drain regions

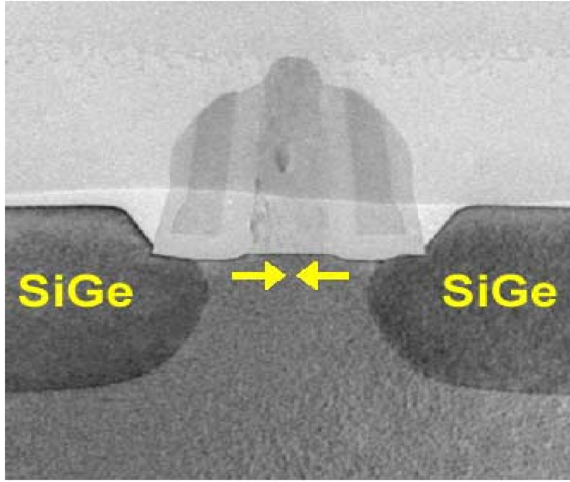



**FIGURE 7.9**
TEM of pMOS Transistor. A strained epitaxial SiGe film is embedded into the source drain region to induce compressive strain in the channel region. After T. Ghani et al., *IEEE IEDM Tech. Dig.*, 2003(978-980). With permission.

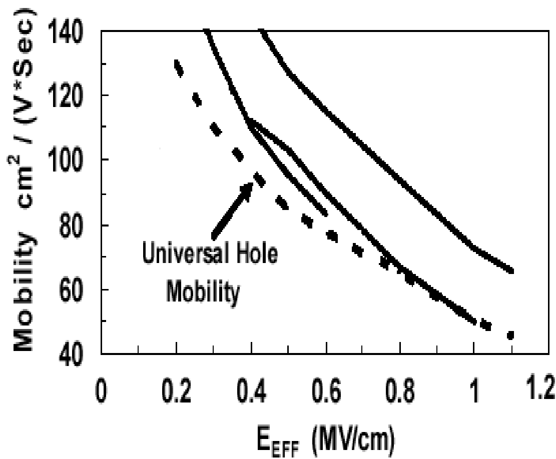



**FIGURE 7.10**
Hole mobility as a function of vertical effective field for uniaxially strained pMOS device relative to conventional biaxially strained substrate techniques. After T. Ghani et al., *IEEE IEDM Tech. Dig.*, 2003(978-980). With permission.

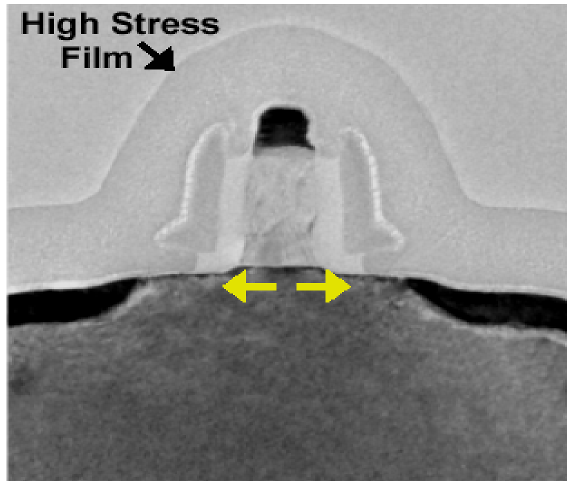



**FIGURE 7.11**
TEM of 45-nm nMOS Transistor. nMOS device is capped with a specially engineered high tensile stress silicon nitride layer to induce tensile channel strain. After T. Ghani et al., *IEEE IEDM Tech. Dig.*, 2003(978-980). With permission.

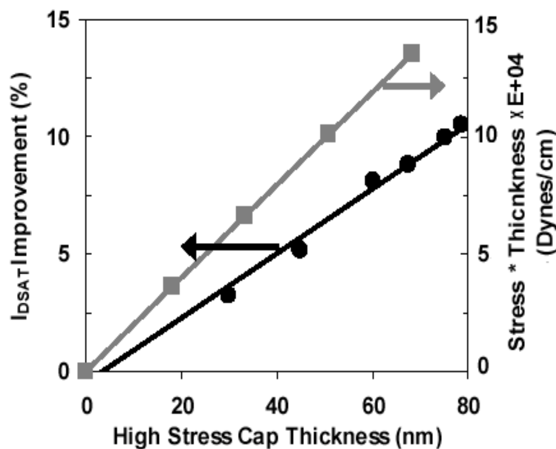



**FIGURE 7.12**
Measured integrated film stress and the resultant nMOS $I_{dsat}$ improvement with increasing tensile film thickness. Tensile film thickness is selected to increase nMOS $I_{dsat}$ by 10% for the devices. After T. Ghani et al., *IEEE IEDM Tech. Dig.*, 2003(978-980). With permission.

to create tensile strain in nMOS channel.

Fig. 7.12 shows modulation of measured integrated film stress and the resultant nMOS $I_{dsat}$ improvement with increasing tensile film thickness for the devices. Capping layer thickness is selected to be $\sim$ 75 nm for the devices to provide 10% nMOS $I_{dsat}$ gain from tensile channel strain. This approach to strain engineering has an advantage over the conventional biaxially strained substrate technique in that it uses standard Si wafers and avoids the cost, defect and other process-integration issues associated with SiGe wafers.

## 7.3 Strained-Si MOSFETs with SiC-Stressor

Recently there has been a great deal of interest in channel engineering through the introduction of local stress. Both the SiGe and SiC have been used in the source/drain regions to introduce stress locally for device drive current enhancement. A novel 50-nm gate length strained-Si n-MOSFET comprising SiC S/D regions, metal gate, and high-k gate dielectric was demonstrated [19, 20]. Fig. 7.13 shows the transistor structure with $Si_{1-y}C_y$ stressors in the source and drain regions. The SiC regions act as stressors, giving rise to lateral tensile strain and vertical compressive strain in the channel to enhance electron mobility. In addition, the SiC/Strained-Si heterojunction at the source end of the transistor provides for enhanced electron injection velocity from the source (Fig. 7.13). The $I_{ds}$-$V_{ds}$ characteristics of a 50-nm gate length $L_g$ transistor with SiC S/D regions shows 50% enhancement in drive current over a control transistor at a gate over-drive ($V_g$-$V_t$) of 1.0 V, as shown in Fig. 7.14.

The spacing $L_g$ between the stressors, the C mole fraction y, h, and d are varied. The theoretical limit of the channel stress is determined by the maximum stress that can be generated at the Si/SiC interface before dislocations are generated. The maximum stress depends directly on the carbon mole fraction and the limit is defined by the following equation $\sigma = E\epsilon/(1\text{-}\nu) = 76.5y$ (GPa) where $E^o$ is silicon Young's modulus, $\nu$ is silicon Poisson's ratio, and $\epsilon$ is the strain induced by the difference of the Si and $Si_{1-y}C_y$ lattice constants ($\epsilon = 0.34y$). When the uniaxial tensile strain is induced in the channel region, an enhanced electron mobility is obtained. Also, the conduction band offset $\Delta E_c$ at the heterojunction between the SiC source and the strained-Si channel enables injection of electrons with additional energy $\Delta E_c$, or at a higher velocity, leading to higher $I_{ds}$.

The partially-relaxed SiC stressors in the source and drain regions tensilely strain the Si channel laterally, leading to a large tensile stress $\sigma_{xx}$ that extends throughout the channel region. It has been observed that in case of anisotropic recess etch, for a given carbon mole fraction the amount of stress in the channel

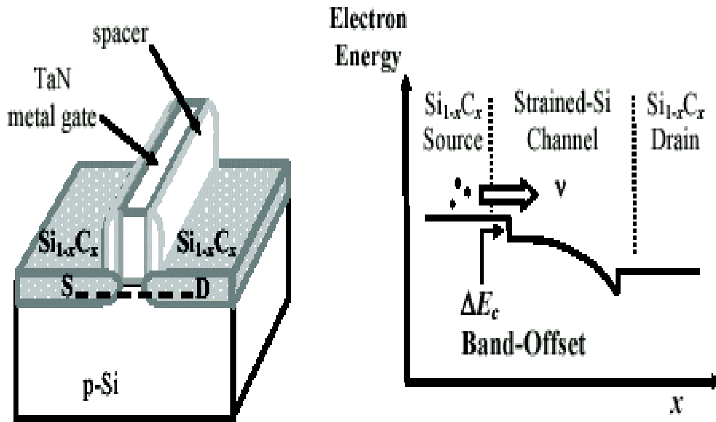



**FIGURE 7.13**
Schematic showing n-MOSFET structure with $Si_{1-x}C_x$ source and drain (S/D) regions. Gate length $L_g$ is 50 nm. Conduction band profile along S-to-D is also shown, illustrating enhanced electron injection velocity from the source into the strained-Si channel. After K. W. Ang et al., *IEEE IEDM Tech. Dig.*, 2004(1069-1071). With permission.

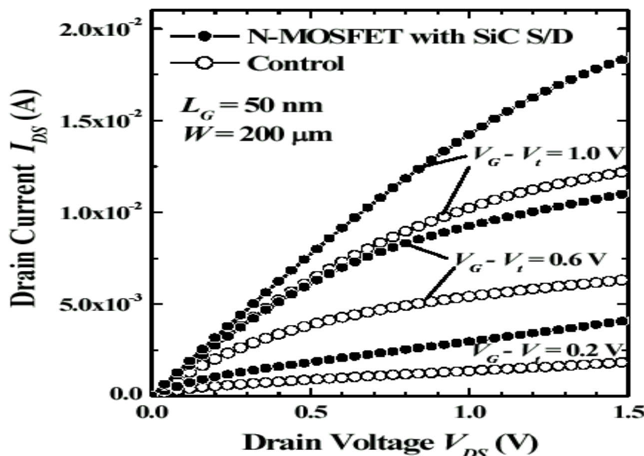



**FIGURE 7.14**
Output characteristics at various gate overdrives ($V_g$-$V_t$). n-MOSFET with SiC S/D regions demonstrate 50% $I_{ds}$ enhancement at a gate length of 50 nm. After K. W. Ang et al., *IEEE IEDM Tech. Dig.*, 2004(1069-1071). With permission.

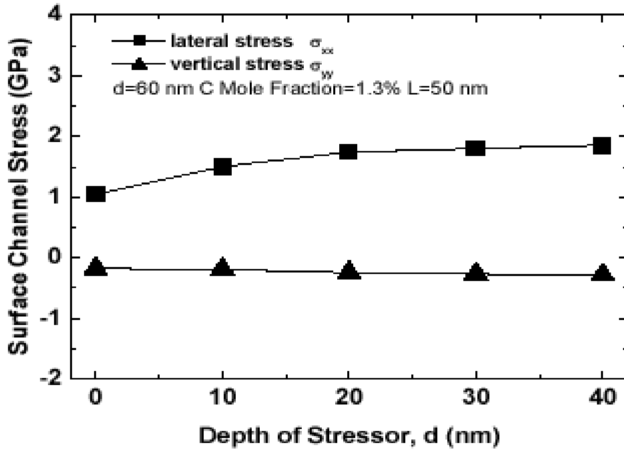



**FIGURE 7.15**
Raising the stressor height h leads to more tensile lateral stress $\sigma_{xx}$ After J. Huang and S. T. Chang, *Semicond. Dev. Res. Symp.*, 2005(107-108). With permission.

is determined by etch depth, which correlates to the SiC thickness, and the etch shape. The $Si_{1-y}C_y$ stressors affect two major strain components, the lateral stress $\sigma_{xx}$ and the vertical stress $\sigma_{yy}$. If the lattice of $Si_{0.987}C_{0.013}$ is fully strained to assume the lattice constant of the underlying Si substrate, then the value of strain in the $Si_{0.987}C_{0.013}$ source/drain region will be 0.45%, and there will be no strain in the Si substrate.

Commercial process simulator FLOOPS-ISE [21] generated results are shown in Figs. 7.15-7.18. The impact of interstressor spacing L, the recess depth d, raised height h, and C content y in the $Si_{1-y}C_y$ stressors on the spatially stressed components in the channel region are shown. Increasing the stressor height h in a raised S/D structure while maintaining the same stressor depth d leads to a more tensile stress $\sigma_{xx}$ (Fig. 7.15). Figure 7.16 shows that for a given L of 50 nm and C content of 1.3%, increasing the depth of the SiC stressor, increases lateral stress $\sigma_{xx}$ in the Si channel while slightly decreasing vertical stress $\sigma_{yy}$ in that one.

For a given L of 50 nm and d = 60 nm, increasing the C mole fraction y, i.e., the lattice mismatch between the SiC stressor and the channel, increases the magnitude of both $\sigma_{xx}$ and $\sigma_{yy}$ linearly (Fig. 7.17). For a given y of 1.3% and d = 60 nm, decreasing L increases the magnitudes of both $\sigma_{xx}$ and $\sigma_{yy}$, as shown in Fig. 7.18. In general, $\sigma_{xx}$ is more uniformly distributed and is larger in magnitude than $\sigma_{yy}$.

The electrical characteristics of the strained-Si n-MOSFETs were simulated with DESSIS [22] using the strain-induced mobility models [23] to account for

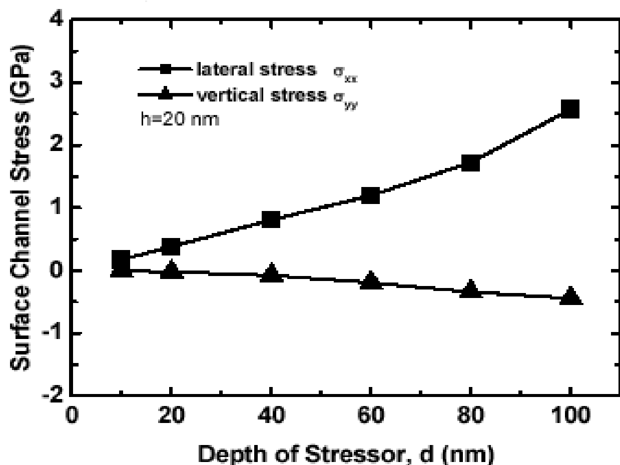



**FIGURE 7.16**
Lateral and vertical strain components, $\sigma_{xx}$ and $\sigma_{yy}$, respectively, in the surface channel plotted as a function of the depth of the $Si_{1-y}C_y$ stressor. After J. Huang and S. T. Chang, *Semicond. Dev. Res. Symp.*, 2005(107-108). With permission.

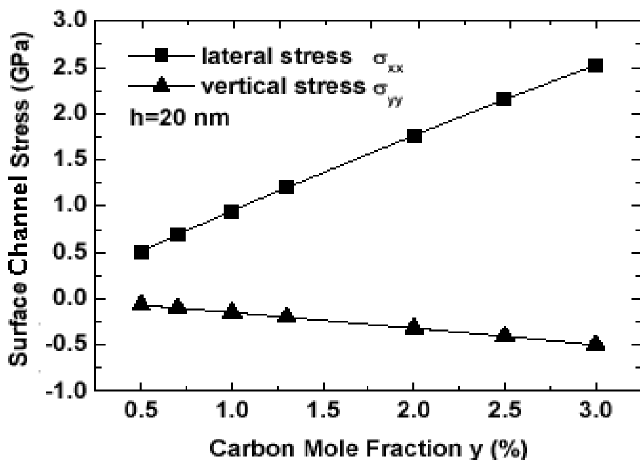



**FIGURE 7.17**
Lateral and vertical strain components, $\sigma_{xx}$ and $\sigma_{yy}$, respectively, in the surface channel plotted as a function of the C mole fraction y in the $Si_{1-y}C_y$ stressor. After J. Huang and S. T. Chang, *Semicond. Dev. Res. Symp.*, 2005(107-108). With permission.

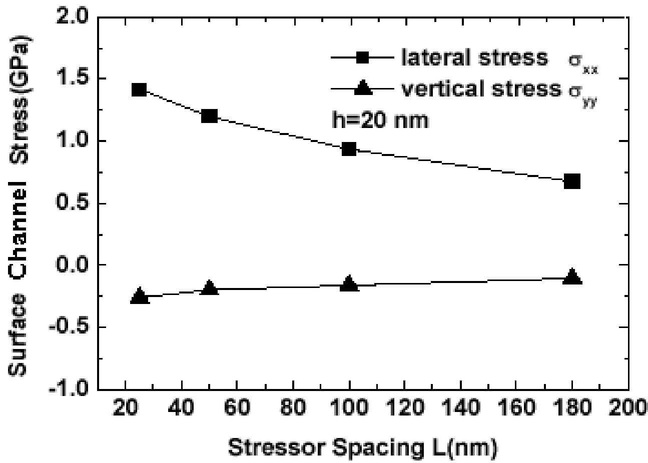



**FIGURE 7.18**
Lateral and vertical strain components, $\sigma_{xx}$ and $\sigma_{yy}$, respectively, in the surface channel plotted as a function of the spacing between the source stressor and drain stressor. After J. Huang and S. T. Chang, *Semicond. Dev. Res. Symp.*, 2005(107-108). With permission.

the change of mobility in highly strained regions. A spatial averaging of the stress components over the region where the inversion electron resides was used to get the mobility enhancement. In general, the contribution of lateral stress $\sigma_{xx}$ to mobility enhancement is found to be larger than that of vertical stress $\sigma_{yy}$. 100% mobility enhancement is achievable with L = 50 nm, d = 60 nm, h = 20 nm, and y = 1.3%, as shown in Fig. 7.19. The thick lines are drawn with L of 5 nm, 50 nm, 100 nm, and 180 nm to guide the eye for the purpose of the optimal design. The mobility enhancement is approximately two times saturation drain current enhancement due to smaller sensitivity of saturation drain current on channel electron mobility (see Fig. 7.19(b)).

In the following, the magnitude and distribution of stress components, the origin of the stress field, and their relationship to electron mobility enhancement have been discussed. It is shown that the strain effect due to the SiC S/D stressors as well as the increased electron injection velocity, may play an important role at sub-100 nm gate lengths. Reducing the interstressor spacing, and increasing the C content and the recessed depth/raised height of the SiC stressors are three ways to achieve high strain levels in the Si channel region for drive current and enhanced electron mobility in n-channel metal-oxide-semiconductor transistors.

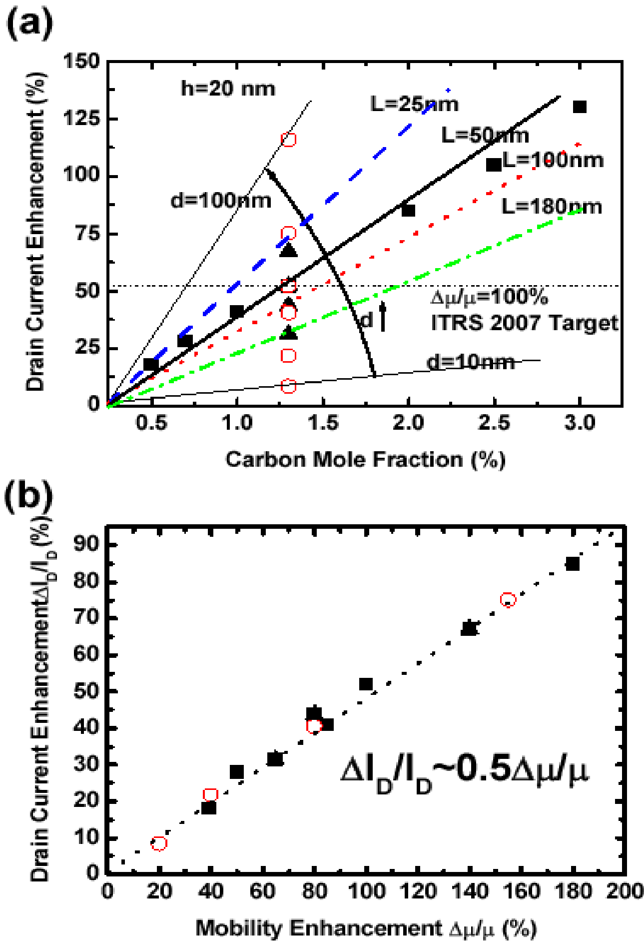



**FIGURE 7.19**

(a) Increasing the Carbon mole fraction in $Si_{1-y}C_y$ stressor and reducing L and larger d leads to larger lateral stress $\sigma_{xx}$ and vertical stress $\sigma_{xx}$. (b) Correlation of the simulated saturation drain current increase to mobility increase. The dotted line is drawn with a slope of 0.5 to guide the eye. The solid squares denote the case of L = 50 nm and d = 60 nm. The solid triangles denote the case of L = 50 nm and carbon mole fraction of 1.3%. The open circles denote the case of L = 50 nm and carbon mole fraction of 1.3%, and h = 20 nm in all cases. After J. Huang and S. T. Chang, *Semicond. Dev. Res. Symp.*, 2005(107-108). With permission.

## 7.4 Biaxial Strain: Process Flow

Strained-Si technology offers opportunities to the CMOS manufacturers to obtain performance and flexibility without increasing costs, and to the circuit designers of products that demand the latest CMOS capabilities at reasonable cost. SiGe-based devices have shown advantageous DC and rf performance using the enhanced electronic properties associated with strain engineering and heterojunction energy barriers. An important point to mention is that the equipment and processes used in standard epitaxial Si wafer processing can be readily implemented for the manufacture of strained-Si substrates. Once wafer costs are addressed, the next consideration for this technology is the ease of fabrication of strained-Si devices and circuits. The typical sequence of steps for the fabrication of CMOS circuits upon strained-Si substrates is unchanged, which means that no equipment changes or specialty processes are required. Overall, the process changes that are necessary can be simpler than those required for moving from one node of standard CMOS technology to the other. However, there are specific modules that need to be updated, as the material properties of the strained-Si substrate are somewhat different. An additional benefit is the fact that no new equipment is required with strained-Si technology.

Perhaps the one feature of strained-Si technology that distinguishes it most from bulk-Si technology is the fact that the thin film of strained-Si, which is limited in thickness due to critical thickness constraints, cannot be stripped indiscriminately. In practice, as long as approximately 5 nm of strained-Si remains at the end of the processing sequence, the transistors will exhibit optimal performance.

With these constraints in mind, strained-Si transistors have been fabricated at aggressive manufacturing nodes. Various groups have reported demonstration of sub-100 nm channel length MOSFETs on strained-Si on relaxed-SiGe heterostructures [14, 24, 25, 26, 27]. In both nMOS and pMOS, the electrostatic design of the strained-Si device is very similar to the conventional Si MOSFETs. So with careful design that takes into account some of the issues that are described in the following sections, short-channel effects that are equivalent to the unstrained-Si devices can be obtained in strained-Si MOSFETs.

Depending on the extent of transistor engineering, the first generation of strained-Si technology has exhibited approximately 30% drive current enhancement for nMOS transistors, and 5-10% drive current enhancement for pMOS transistors incorporating into existing mask sets and designs, prolonging the life cycle of circuit designs. The infrastructure for design does not change with strained-Si technology, so circuit designers can expect SPICE models and design libraries similar to those they receive from the fabrication.

In the following, we discuss the integration scheme of dual-strained surface-

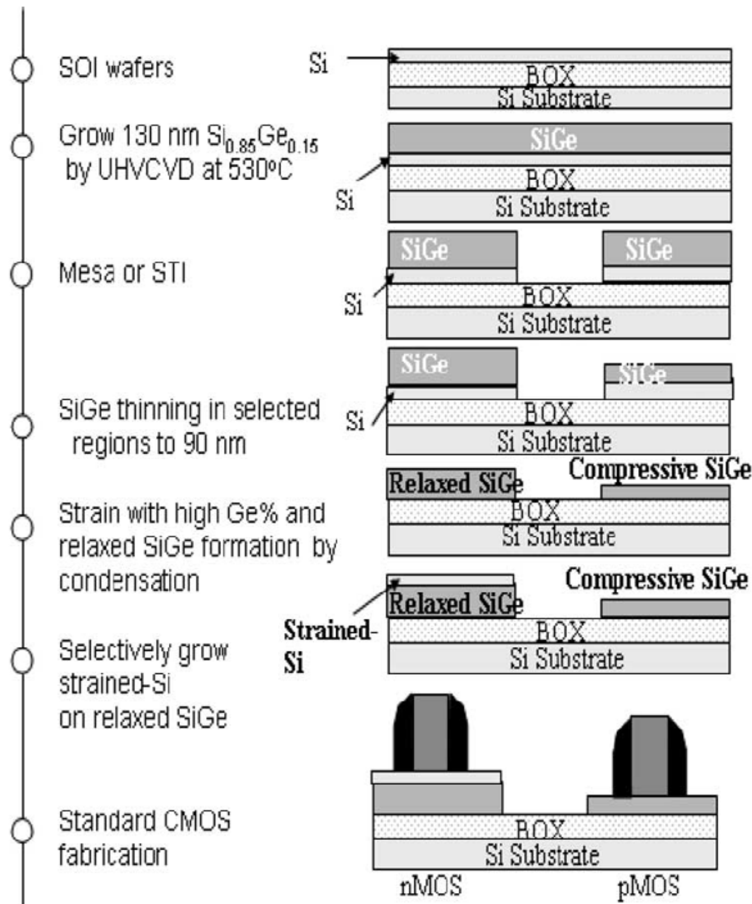



**FIGURE 7.20**
Process flow for fabricating strained and relaxed-SiGe, and the CMOS structure fabricated. After L. K. Bera et al., *IEEE Electron Dev. Lett.*, Vol. 27, 2006(350-353). With permission.

channel CMOS structure, i.e., tensile-strained Si n-MOSFET and compressive strained-SiGe p-MOSFET. This has been accomplished by forming the relaxed and compressive strained-SiGe layers simultaneously on an Si/SiGe-on-insulator (SOI) substrate, through varying SiGe film thicknesses, followed by a thermal condensation technique to convert the Si body into SiGe with different Ge concentration and with different strains (including the relaxed state). A p-MOSFET with compressive ($\epsilon \sim$ -1.07%) strained-SiGe and the n-MOSFET in tensile strained-Si over relaxed-$Si_{0.8}0Ge_{0.20}$ exhibited significant hole (enhancement factor $\sim$ 1.9) and electron (enhancement factor $\sim$ 1.6) mobility enhancements over the Si reference.

Fig. 7.20 describes the process flow and device structures. Starting wafers were n-type (100) 8-in SOI with Si thickness of 75 nm over buried oxide (BOX). Si was thinned to 30 nm using thermal oxidation and wet etch, followed by ultrahigh vacuum chemical vapor deposition of $Si_{0.85}Ge_{0.15} \sim$ 130 nm. This SiGe thickness was kept below the critical thickness for a deposition temperature of 580$^o$C to minimize misfit dislocations. Mesa structures were patterned and etch-stopped on the BOX. Subsequently, two different SiGe thicknesses were obtained on the wafer by selectively masking the location that was intended to retain a thick SiGe ($\sim$ 130 nm) and to etch the unmasked SiGe to $\sim$ 90 nm using HBr/$Cl_2$ chemistry. After photoresist and polymer removal, wafers were subjected to a condensation process started with 1050$^o$C $N_2$ anneal followed by cyclic oxidation and anneal steps at the same temperature to condense Ge into the underlying Si to form Ge-rich SiGe. The $N_2$ annealing is beneficial to homogenize Ge across the SiGe and underlying Si.

After preparing the final SiGe substrates, the oxide was striped selectively from relaxed-SiGe regions using dilute hydrofluoric acid (HF:$H_2O_2$ = 1:25) solution where nMOS fabrication was intended. A thin (15 nm) Si was then epitaxially grown at 580$^o$C on a relaxed-SiGe island selectively to achieve the tensile strained-Si, followed by oxide strip from the rest of the wafer (i.e., compressive strained-SiGe areas). All wafers received DHF (HF:$H_2O_2$ = 1:200) clean, followed by low-temperature gate oxidation by RTO ($t_{ox} \sim$ 2.5 nm on strained-Si and control Si and $\sim$ 4 nm on SiGe as measured by an ellipsometer). The interface qualities for both gate stacks on Si and SiGe channels were found to be similar based on C-V measurement. Interface trap densities were estimated to be $\sim 4\times10^{10}$ cm$^2$ eV$^{-1}$ at midgap for all three cases. Poly-Si $\sim$ 130 nm was deposited at 580$^o$C and patterned. The lightly doped source/drain (S/D) was implanted with As ($8\times10^{14}$ cm$^{-2}$/10 keV) and $BF_2$ ($4\times10^{14}$ cm$^{-2}$/7 keV) for nMOS and p-MOS, respectively. A spacer with low-pressure CVD SiN $\sim$ 800 Å was formed, followed by deep S/D implant and spike anneal (1000$^o$C/$N_2$). After contact, Al metallization sintering was done in forming gas at 420$^o$C/30 min.

Fig. 7.21 shows the output characteristics $I_d$-$V_d$ of n/p-MOS for compressive strained-SiGe and tensile strained-Si in comparison with Si control reference devices. nMOS shows $\sim$ 1.64 drive current enhancement; meanwhile, a significant enhancement of $\sim$ 1.9 is also observed in p-MOS. For both n-

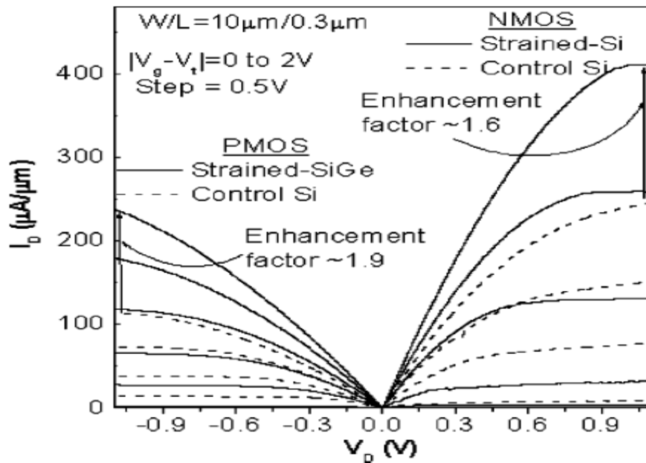



**FIGURE 7.21**
Output characteristics ($I_d$-$V_d$) of n- and p-MOSFETs with strained-Si over relaxed-$Si_{1-y}Ge_y$ (y = 20%) and in pMOS with compressive strained-$Si_{1-z}Ge_z$ (z = 45%), respectively, in comparison with the unstrained-Si reference. The gate drive $V_g$ in all presented has been normalized to $V_t \cdot W$ = 10$\mu$m and $L_g$ = 0.3 $\mu$m. After L. K. Bera et al., *IEEE Electron Dev. Lett.*, Vol. 27, 2006(350-353). With permission.

and p-MOSFETs, the drive current enhancements are consistent with their respective tensile and compressive strain levels. Fig. 7.22 shows the electron and hole mobility enhancement factors obtained as a function of the incorporated Ge mole fraction, along with the published theoretical and experimental data. The mobility performances from control devices are similar to the universal curves for electron and hole. Electron mobility shows an enhancement factor of $\sim$ 1.6 ($\mu_{eff,strained-Si}/\mu_{eff,control}$ = 524/320 at 0.8 MV/cm) over control-Si reference, close to the calculated value. A significant hole mobility improvement is also observed. The enhancement factor of hole mobility is $\sim$ 1.9 ($\mu_{eff,strained-Si}/\mu_{eff,control}$ = 130/67 at 0.8 MV/cm), which is lower than the calculated value but quite close to the bulk strained-SiGe.

High strain and high Ge content increase the difficulty of various process integration issues, and due to the small hole mobility enhancement at high vertical fields, strained-Si is expected to provide only limited enhancement in the future technology nodes. As discussed in chapter 3, strained-Si/relaxed-SiGe heterostructures can be formed on a layer of insulator. Such structures, which are known as SiGe-on-insulator (SGOI), combine the enhanced carrier transport in strained-Si with the advantages of SOI MOSFETs. One major difference between SGOI devices and bulk strained-Si devices on graded buffer is that the SiGe layer in SGOI needs to be significantly thinner.

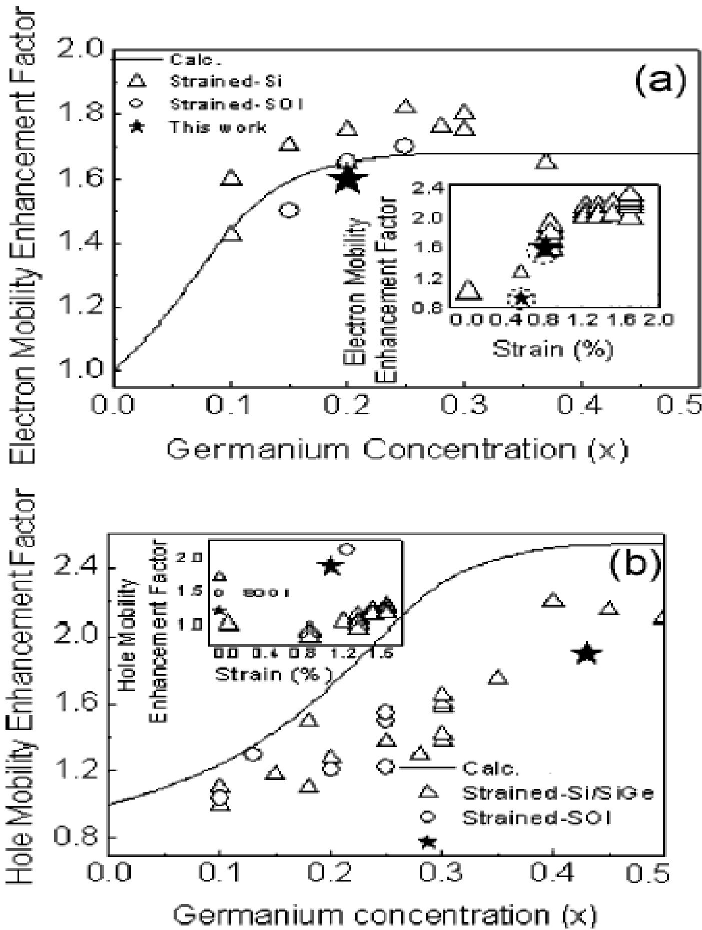



**FIGURE 7.22**

(a) Electron and (b) hole mobility enhancement factor in n-MOSFETs with strained-Si over relaxed-$Si_{1-y}Ge_y$ (y = 20%) and in pMOS with compressive strained-$Si_{1-z}Ge_z$ (z = 45%), respectively, in comparison with calculated and experimental published data. After L. K. Bera et al., *IEEE Electron Dev. Lett.*, Vol. 27, 2006(350-353). With permission.

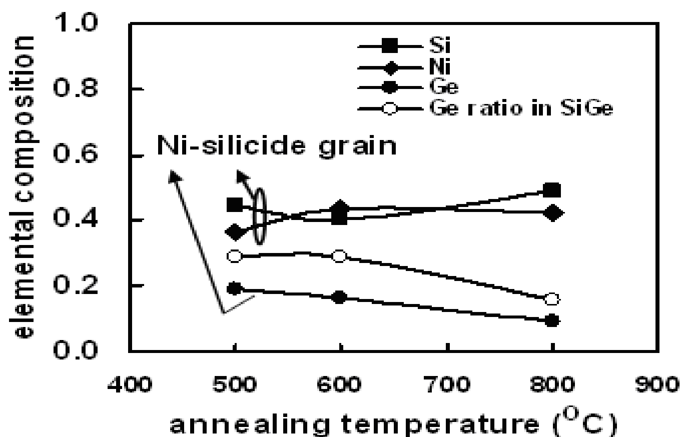



**FIGURE 7.23**
Plots for relative compositions of Ni, Si, and Ge in the nickel-germanosilicide grain. Ge is gradually expelled and diffused from the silicide grains with increasing annealing temperature. After S. Chattopadhyay et al., *Solid-State Electron.*, Vol. 48, 2004(1407-1416). With permission.

Reduction of the process thermal budget may lead to an unacceptable increase in device parasitics, compromising the advantages gained by the use of high-mobility channel. A further issue to be considered is the consumption of the strained-Si surface layer during device fabrication through different cleaning and oxidation. To meet the requirement for desired performance, sufficient strained-Si must be left for the channel formation after processing. The strained-Si channel thickness is a crucial parameter in this complex design space. If the channel is too thin, Ge diffusion up to the Si/$SiO_2$ interface may give rise to an unacceptable level of interface charge. Moreover, the strained-Si channel will be unable to contain the inversion layer charge, leading to an increase in carrier scattering and degraded device performance.

The implementation of a reliable and reproducible SALICIDE (self-aligned silicide) process in strained-Si/$Si_{1-x}Ge_x$ $Si_{1-x}Ge_x$ CMOS technology is a major challenge. The critical thickness for strained-Si is very small ($<$ 10 nm) for 20-30% of Ge composition in the virtual substrate [28]. The remaining strained-Si layer in the source and drain regions after undergoing different cleans is too thin for silicidation. As the strained-Si layer in the source/drain regions is thin, the thermal reaction during silicidation will certainly extend to the underlying $Si_{1-x}Ge_x$ layer and $Si_{1-x}Ge_x$ will take part in the thermal reaction. The presence of Ge in the $Si_{1-x}Ge_x$ layer changes the silicide reaction mechanism in many ways. For example, $Si_{1-x}Ge_x$ inhibits the transition of cobalt silicide to its lowest resistive disilicide ($CoSi_2$) phase [29]. Also, Ge

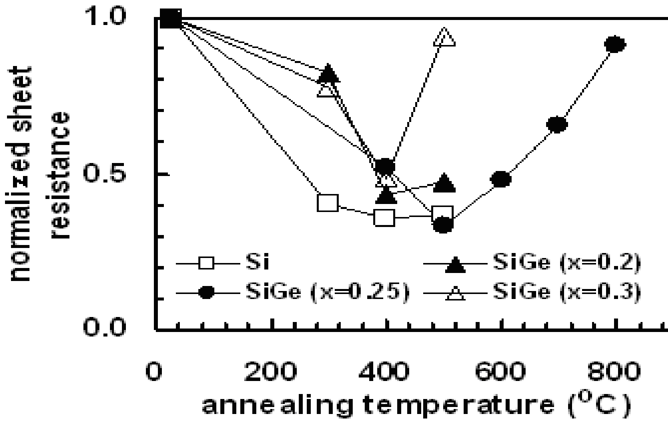



**FIGURE 7.24**
Plots of normalized sheet resistance of the Ni-silicide films annealed at different temperatures for 60 s. Plots show that the films achieve their lowest resistive phase in the range of 400-600$^o$C. After S. Chattopadhyay et al., *Solid-State Electron.*, Vol. 48, 2004(1407-1416). With permission.

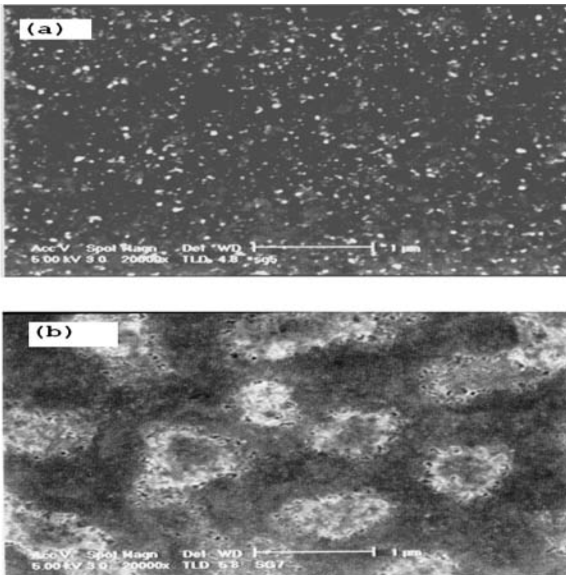



**FIGURE 7.25**
Scanning electron microscope (SEM) images for the surface of Ni-silicided films annealed at (a) 500$^o$C and (b) 800$^o$C for 60 s. After S. Chattopadhyay et al., *Solid-State Electron.*, Vol. 48, 2004(1407-1416). With permission.

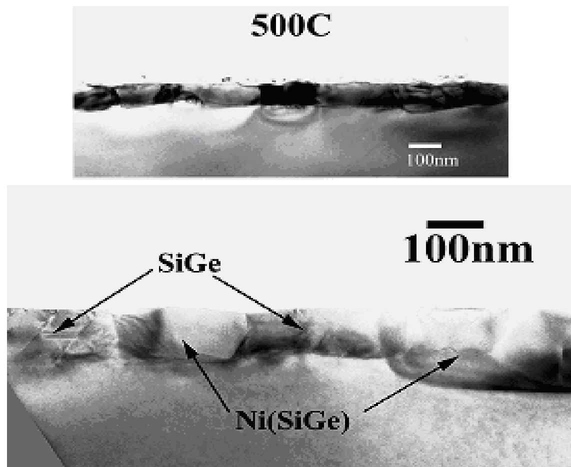



**FIGURE 7.26**
Transmission electron microscope (TEM) images for the surface of Ni-silicided films annealed at (a) 500$^o$C and (b) 700$^o$C for 60 s. After S. Chattopadhyay et al., *Solid-State Electron.*, Vol. 48, 2004(1407-1416). With permission.

promotes agglomeration during silicidation. Using an elevated source/drain is a solution [30] but its implementation needs extra cleaning and processes, such as selective epitaxy.

$TiSi_2$ is the most commonly used silicide for SALICIDE technology in conventional Si MOSFETs. But its transition to its lowest resistive $TiSi_2$ phase (C-54) requires a high annealing temperature ($> 850^o$C). There are increased risks of strain relaxation and strong agglomeration at such high temperatures. Rim et al. [31] used Ti-SALICIDE process at low thermal budget for the fabrication of strained-Si MOSFET and the resulting resistivity of the silicide film was high. Nickel silicide (NiSi) is a potential candidate for deep submicron strained-Si/SiGe CMOS devices because of its low Si consumption, low thermal budget, and one-step annealing to achieve its lowest resistive phase [32]. Moreover, Ni-silicide can support a low resistant polycide line down to $< 50$ nm.

When an Ni film is deposited on SiGe layer and annealed, Ni-germanosilicide $[Ni_y(Si_{1-x}Ge_x)_{1-y}]$ is formed [17, 33]. The reaction mechanism is complex compared with the reaction with pure Si. The heats of formation for NiSi and NiGe are -45 and -32 kJ/mol, respectively. Thus, thermodynamically Ni prefers to react with Si and as a result, Ge is repelled and diffused from the germanosilicide grains and segregates inbetween the silicide grains. Fig. 7.23 shows the relative composition of the elements Si, Ni, and Ge in the germanosilicide grain when rapid thermally annealed at different temperatures for 60 s. The Ge content in germanosilicide grains decreases with increasing annealing temperature.

Plots of the normalized sheet resistance are shown in Fig. 7.24. The silicide films were grown on relaxed-$Si_{1-x}Ge_x$ layers of different Ge content and annealed at different temperatures for a systematic study. Results for a Ni-silicide film grown on a Si substrate are also included. All the silicide films achieve the lowest sheet resistance value in the temperature range of 400-600$^o$C. The plots indicate that the process window for Ni-silicide film on $Si_{1-x}Ge_x$ varies with Ge composition in the alloy.

Figs. 7.25(a) and (b) are the scanning electron microscope (SEM) images which show Ni-germanosilicide films annealed at 500 and 800$^o$C, respectively. Up to an annealing temperature of 500$^o$C, the film is smooth and continuous but as the annealing temperature increases the film loses its uniformity and transforms into discrete silicide islands distributed over the $Si_{1-x}Ge_x$ layer due to strong agglomeration.

Figs. 7.26(a) and (b) show transmission electron microscope (TEM) images for Ni-silicided films annealed at 500 and 700$^o$C, respectively. The silicide/$Si_{1-x}Ge_x$ interface is apparently very smooth at 500$^o$C. But the 700$^o$C annealed film shows that some $Si_{1-m}Ge_m$ (m $\neq$ x) grains have started to appear inbetween the silicide grains, making the film discrete. Further studies revealed that at higher annealing temperatures more Ge diffuses out from the germanosilicide grains and reacts with substrate $Si_{1-x}Ge_x$ layer, creating higher Ge content $Si_{1-m}Ge_m$ grains [17, 33]. Also, it is reported that with increasing annealing temperature, germanosilicide grains migrate into the $Si_{1-x}Ge_x$ virtual substrate, increasing Ge concentration on the top $Si_{1-x}Ge_x$ layer [34].

Generally, three areas of development that need to be considered when considering the commercial use of global strained-Si technology:

- Concurrent optimization of nMOS and pMOS transistors needs to be explored more thoroughly. A systematic development effort is required to elucidate the complete benefits of global strained-Si technology.
- Defects (misfit dislocations) at gross levels at the interface between strained-Si and the $Si_{1-x}Ge_x$ alloy have been shown to increase transistor leakage levels. However, no significant results have been reported on the mechanisms for the leakage or methods to mitigate such leakage. Further research and development is required to improve the understanding on this topic.
- Leakage and reliability mechanisms associated with threading dislocations and other defects are not yet clearly understood.

## 7.5 Scaling of Strained-Si MOSFETs

Mizuno et al. [35] have reported the physical mechanisms for both the electron and the hole mobility reductions in thin strained-Si structures on strained-

SOI. From low- and room-temperature measurements of the electron mobility in thin and thick strained-Si structures, the decrease in the electron mobility at a low transverse electric field has been attributed to the enhanced phonon scattering by the quantum mechanical confinement of inversion electrons in the conduction band well of the strained-Si layer. On the other hand, the reduction in hole mobility at higher $E_{eff}$ is attributable to the reduced energy-splitting between the light and the heavy hole bands in the inversion layers, because a smaller subband energy shift of the inversion heavy holes diminishes the energy splitting introduced by the strain in Si layers with increasing $E_{eff}$. However, the reduction in hole mobility in the thinner strained-Si layers at lower $E_{eff}$ is caused by Coulomb scattering due to a higher interface state density in p-MOSFETs, which can be associated with the $GeO_2$ formation in the gate oxide layer generated by Ge atoms diffused from SiGe layers. In the following, we discuss the device design for FD strained-SOI CMOS under sub-100 nm regimes for high-end logic ULSIs, considering the physical mechanism for the mobility reduction. As a result of these mobility-lowering mechanisms, a new device-design guideline regarding the strained-Si thickness and the Ge content is needed to realize high-speed thin film strained-SOI CMOS devices in 25-nm gate length regimes.

We discuss below the design for the strained-Si thickness and the Ge content for realizing the carrier mobility enhancement even at high $E_{eff}$ which greatly affects the drain current. SCEs need to be suppressed by thinning the strained-Si-SiGe body for the reduction in the increase of the Coulomb scattering associated with the higher channel dopant. Therefore, the total thickness of the strained-Si and the SiGe $T_{sc}$ layers should be thinner than $L_{eff}/4$ to suppress SCE, according to the empirical knowledge in FD SOI-MOSFETs, which is also verified by the device simulation of the strained-SOIs [36]. As a result, when $L_{eff}$ is 25 nm, $T_{sc}$ needs to be smaller than 6.3 nm. According to the above discussion, it is necessary to satisfy the following factors:

- the lower limit of the Ge content;
- the upper limit of strained-SOI thickness based on the SCE;
- the lower limit of $T_{sc}$ to suppress the increase in both the phonon scattering due to the confinement effects of electrons (see Fig. 7.27) and the Coulomb scattering caused by the holes (see Fig. 7.28), and
- the upper limit originating in the critical thickness of the strained-Si on the relaxed-SiGe layers to avoid the relaxation of the strain in the strained-Si layer [28].

In the case of p-MOS, a minimum Ge mole fraction of 0.3 is needed because the energy splitting between LH and HH should be larger than the optical phonon energy even at high $E_{eff}$. On the other hand, Ge mole fraction for nMOS should be higher than 0.1, according to the relationship between the electron mobility enhancement and the Ge content [37]. In the case of nMOS, Fig. 7.27 shows that $T_{Si}$ of 3 nm is not the limitation, because there is no degradation of electron mobility even at high $E_{eff}$. However, as observed from Fig. 7.28, the hole mobility degradation at $T_{Si}$ of 3 nm, therefore, $T_{Si}$

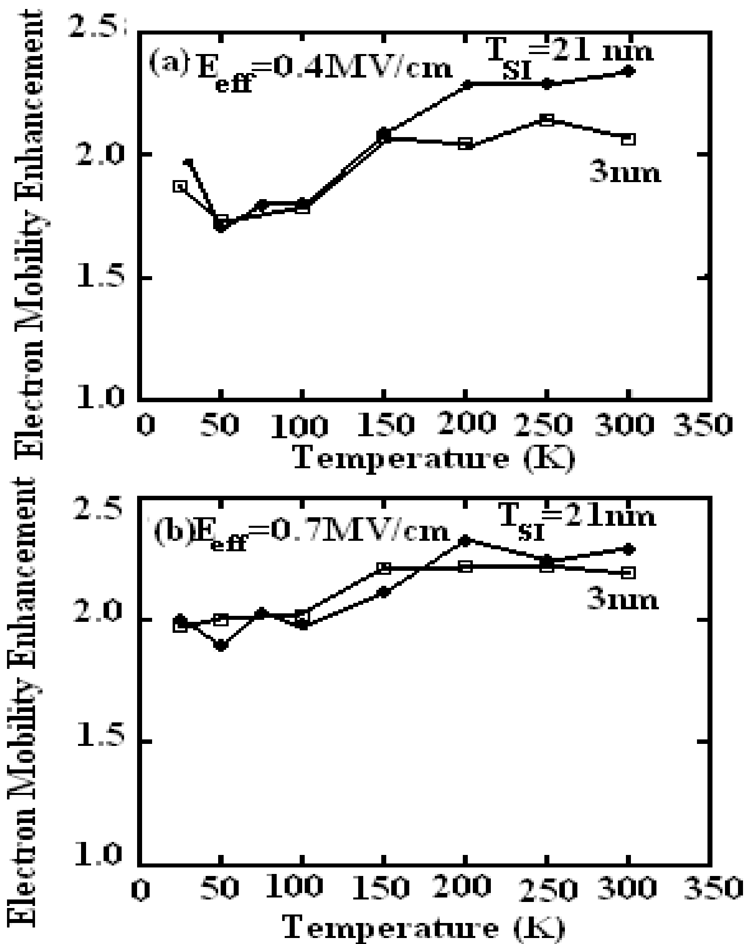



**FIGURE 7.27**
Electron mobility enhancement against that of control SOI MOSFETs vs. temperature at (a) low (0.4 MV/cm) and (b) high electric field (0.7 MV/cm). The circles and the squares show the results of $T_{Si}$ of 21 nm and 3 nm, respectively. (b) Even at low temperatures and a high electric field, the electron mobility is still enhanced. After T. Mizuno et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(1114-1121). With permission.

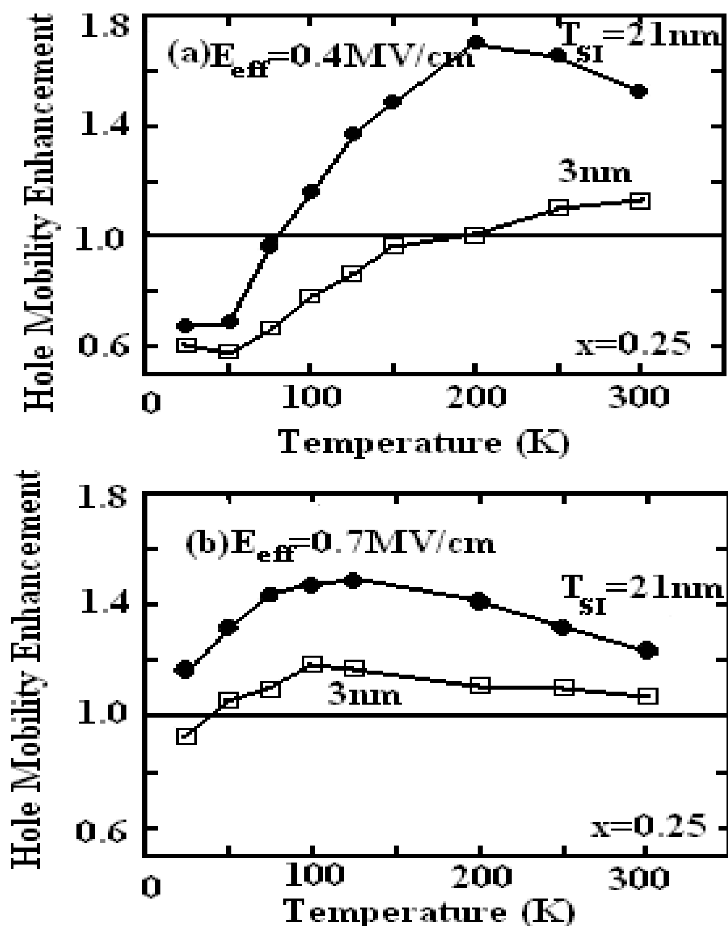



**FIGURE 7.28**
Hole mobility enhancement of $T_{Si}$ of 21 nm (circles) and 3 nm (squares) against that of control SOI MOSFETs vs. temperature at (a) low (0.4 MV/cm) and (b) high electric fields (0.7 MV/cm). The hole mobility enhancements peak at around 150 K. The hole mobility enhancements at low $E_{eff}$ rapidly decrease with temperature lower than 200 K. After T. Mizuno et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(1114-1121). With permission.

needs to be thicker than at least 3 nm for p-MOSFETs. Fig. 7.29 summarizes the conditions for the Ge content and to satisfy the above three factors. The optimum conditions for the Ge content, as in the dashed area in Fig. 7.29, are obtained where $x \geq 0.3$ and $3 \text{ nm} \leq T_{Si} \leq 6.3 \text{ nm}$, when the channel length is 25 nm.

Strained-Si channel MOSFETs have been shown to be very promising for CMOS applications beyond the 50-nm technology node since in strained-Si layers, grown on relaxed-SiGe, carrier mobility increases. However, recent studies show that the mobility of strained-Si n-MOSFETs degrades as the gate length becomes shorter [31], which raises scalability issues in strained-Si MOSFETs. Stress relaxation in strained-Si MOSFETs can be significant in the presence of compressive stress imposed by trench isolation, especially for highly scaled active regions. Stress of the strained region is reduced by 2/3 when the active region is scaled from $L_{active} = 0.4\ \mu\text{m}$ to $0.1\ \mu\text{m}$. Mobility can be lower by 50% for narrow active widths resulting from the strain relaxation. The strain relaxation may restrict the use of strained-Si MOSFETs for technology nodes beyond 25 nm.

Also, the thermal conductivity of $Si_{0.8}Ge_{0.2}$ is 15 times lower than that of bulk-Si, such that self-heating problems similar to that for SOI can cause performance degradation in strained-Si devices. This imposes technology limits of strained-Si devices in terms of electrical and thermal characteristics. Electrothermal characteristics of strained-Si MOSFETs, operating in the high-current regime, have been studied using device simulation [38]. Shallow trench isolation-induced compressive stress reduces carrier mobility and drive current in bulk nMOS. The compressive stress in silicon from trench isolation makes the nMOS drive current highly sensitive to the transistor layout. Thus, problems arise when the process-induced stress relaxes the tensile stress in the strained layer for scaled active regions. Choi et al. [38] have performed detailed stress analysis based on lattice and thermal mismatch models for SiGe and silicon depositions, as well as STI formation by using 2D and 3D process/device simulations. Fig. 7.30 shows the cross-section of a strained-Si layer, which is grown on the relaxed-SiGe seed layer with a Ge mole fraction of 0.2. The strained-Si nMOS has a strained-Si layer of 15 nm thickness and source/drain junction depth of 50 nm. The gate length is 130 nm and the gate oxide thickness is 2.5 nm for both structures. The channel dopant concentration for both structures is about $1\times10^{18}\ \text{cm}^{-3}$ in the presence of halo implantation. The lattice constant of SiGe is based on Vegard's law and the lattice mismatch strain is calculated when Si is grown on the SiGe layer. Stress relaxation is calculated for a structure where the trench is etched out and oxide filling is then used.

Fig. 7.31 shows two-dimensional stress analysis results of a strained active region with STI for different lengths of active layers ($L_{active}$). The initial stress ($\sigma_{xx}$) in the strained Si layer is 1400 MPa for $L_{active} = 0.4\ \mu\text{m}$ ($L_g = 0.1\ \mu\text{m}$), which is reduced to 1033 Mpa after the STI process. The tensile stress in the strained layer is reduced down to 600 MPa and 350 MPa as $L_{active}$ is

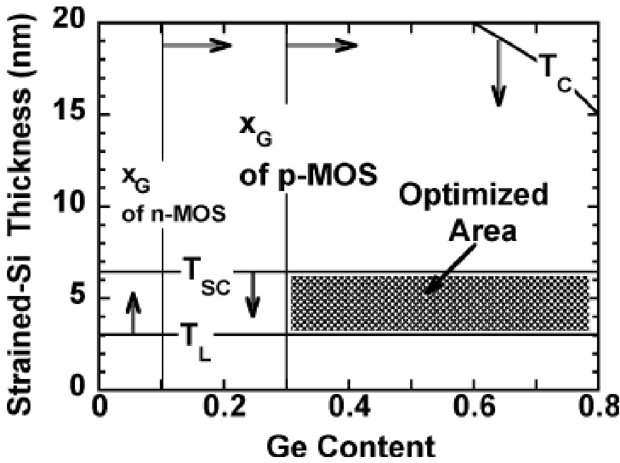



**FIGURE 7.29**
Design guideline for $T_{Si}$ and Ge content in 25-nm gate length region. $T_{Sc}$, $T_l$, $T_c$, and x are the upper limit thickness of strained-SOIs for suppressing the SCEs, the lower limitation of $T_{Si}$ for quantum mechanical confinement, the critical thickness of strained-Si on SiGe layers, and the lower limit of the Ge content for preventing the reduction of the energy splitting between the heavy and the light holes at higher $E_{eff}$, respectively. The dashed region shows the optimized conditions for $T_{Si}$ and the Ge content. After T. Mizuno et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(1114-1121). With permission.

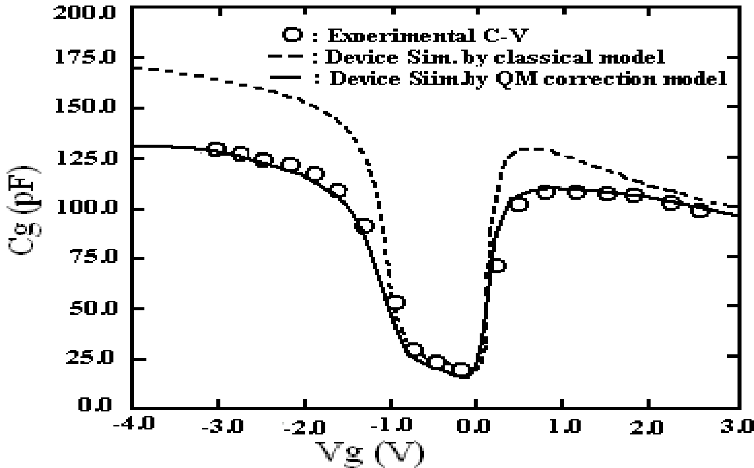



**FIGURE 7.30**
A cross-section of strained-Si active layer on $Si_{0.8}Ge_{0.2}$ layer. After R. W. Dutton and C.-H. Choi, Presentation in NSTI Nanotechnology Conference 2004.

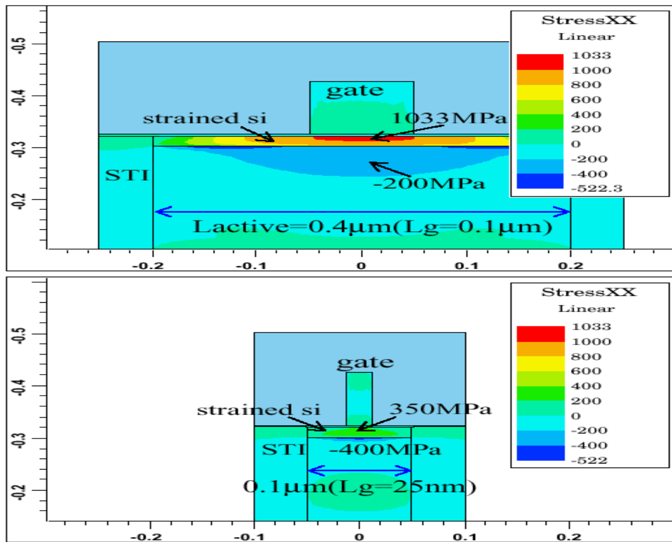
scaled to $L_{active} = 0.2$ $\mu$m ($L_g = 50$ nm) and $L_{active} = 0.1$ $\mu$m ($L_g = 25$ nm), respectively. Strain relaxation is dependent not only on the device geometry, but also on thermal annealing. Three-dimensional device simulations are performed based on the stress-induced bandgap change model, in which the bandgap varies with regard to mechanical stress and strain tensors in silicon region.

Fig. 7.32 shows simulated $I_d$-$V_g$ and $g_m$-$V_g$ based on stress calculations for different annealing conditions after STI process: 1 s, 800$^o$C and 10 s, 850$^o$C. This implies that mobility enhancement degrades for the higher temperature annealing condition that enhances strain relaxation. Fig. 7.33(a) shows $\sigma_{xx}$ contours for different active widths; stress relaxation in the strained layer is enhanced for the narrower active width due to impact of the surrounding STI. As a result, mobility enhancement of the narrow width device (i.e., W = 0.2 $\mu$m) is reduced by more than 50% compared to that of the wider device (i.e., W = 1.7 $\mu$m), as shown in Fig. 7.33(b).

Fig. 7.34(a) shows simulated $I_d$-$V_d$ curves for strained- and bulk-Si devices ($V_g = 0$ V). The snapback voltage of the strained-Si device is lower than that for bulk-Si device. The hold voltage ($V_h$), i.e., the minimum voltage required for the bipolar operation, is also lower for the strained-Si device. The second breakdown triggering current ($It_2$) for strained-Si device is higher relative to bulk-Si device. As a result, the power density and peak lattice temperature during the parasitic bipolar operation are lower for the strained-Si device than that for bulk-Si device. The thermal conductivity of $Si_{0.8}Ge_{0.2}$ is 15 times lower than that of bulk-Si, such that self-heating problems similar to that for SOI can cause performance degradation in strained-Si devices.

Fig. 7.34(b) shows peak temperature vs. drain current ($I_d$) for bulk- and strained-Si devices. There is an abrupt increase of lattice temperature at 8 mA/$\mu$m of drain current for the bulk-Si device. as a result, thermal failure occurs for the bulk-Si device ($T > T_c$ 1650$^o$K) when the drain current is 10 mA/$\mu$m ($It_2$) due to its high power density. This implies that even though the thermal dissipation of strained-Si is worse compared to silicon due to the 30 times lower thermal conductivity of SiGe layer, the local temperature overheating can effectively be suppressed in strained-Si devices owing to the higher bipolar gain and current uniformity.

The scaling trend of strained-Si MOSFETs has also been studied via simulation. The 2D process-to-device simulations have been carried out with Silvaco. The DC and AC performance of deep submicron strained-Si devices have been simulated using several transport models viz., extended drift-diffusion and hydrodynamic/NEB models to investigate the technology scaling issues via process simulation. The parameter models and mobility expressions include the effects of alloy composition and mechanical strain on the band structure of Si/SiGe alloys and the corresponding interfaces. Generally, band offsets enable carrier confinement whereas band splitting affects the effective mass and the density of states, reducing the amount of intervalley scattering and, as a result, enhancing carrier transport in the strained layer.



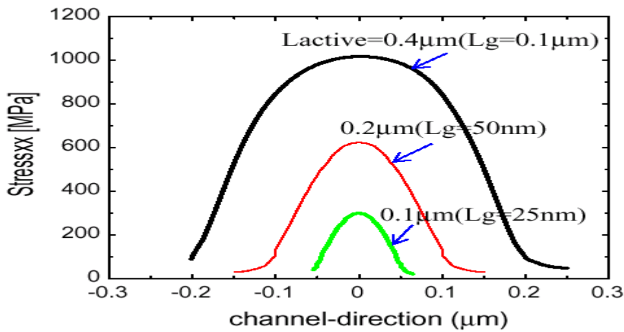




**FIGURE 7.31**

Two-dimensional stress analysis in a strained-Si channel MOSFET with STI. (a) stress ($\sigma_{xx}$) contours for $L_{active} = 0.4$ $\mu$m ($L_g = 0.1$ $\mu$m) and $L_{active} = 0.1$ $\mu$m ($L_g = 25$ nm) and (b) comparison of tensile stresses in strained-Si layer along the channel direction for different active channel lengths. After R. W. Dutton and C.-H. Choi, Presentation in NSTI Nanotechnology Conference 2004.

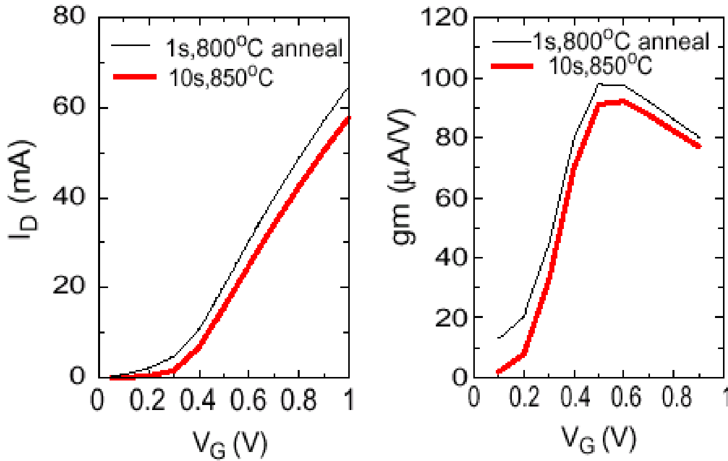



**FIGURE 7.32**
Simulated $I_d$-$V_g$ and $g_m$-$V_g$ based on stress calculation for different annealing conditions after STI process: 1 s, 800$^o$C and 10 s, 850$^o$C. After R. W. Dutton and C.-H. Choi, Presentation in NSTI Nanotechnology Conference 2004.

To predict the performance enhancement via simulation, it is essential to model the band alignment and strain effects within the heterostructure and, as has been discussed before, were used in the simulation. The ATHENA process simulation framework from Silvaco was used for the process simulation of a typical strained-Si MOSFET with scale-down structures. Typical Si CMOS fabrication process steps were used in simulation using SSuprem-4 process simulation suit of ATHENA. Simulation for implantation was performed with $As^+$ and $BF_2{}^+$ implant model, to create deep and shallow S/D regions and the p-well, respectively, for strained-Si. The ion energies were chosen from standard processing data; to locate the peak-implanted concentration near the SiGe/Si heterointerface and the ion dose used was 2-5$\times 10^{15}$ cm$^{-2}$. The simulated device structure is shown in Fig. 7.35.

An important figure-of-merit of RF MOSFETs or HBTs is the cutoff frequency $f_t$. This is the frequency at which the small-signal current gain of the transistor rolls off to unity (i.e., 0 dB). Fig. 7.36 shows the simulated current gain vs. frequency with variable gate length of n-channel MOSFETs. For sub-45 nm gate length n-MOSFETs, the cutoff frequency can reach as high as 210 GHz. Fig. 7.37 shows the minimum noise figure and propagation delay, obtained from the calibrated n-MOSFET from simulation. It is clear that the noise and delay performances degrade with scaling, though the DC performance of the devices is enhanced. Thus, challenges remain regarding the modeling and optimization of strained-Si for RF wireless applications.

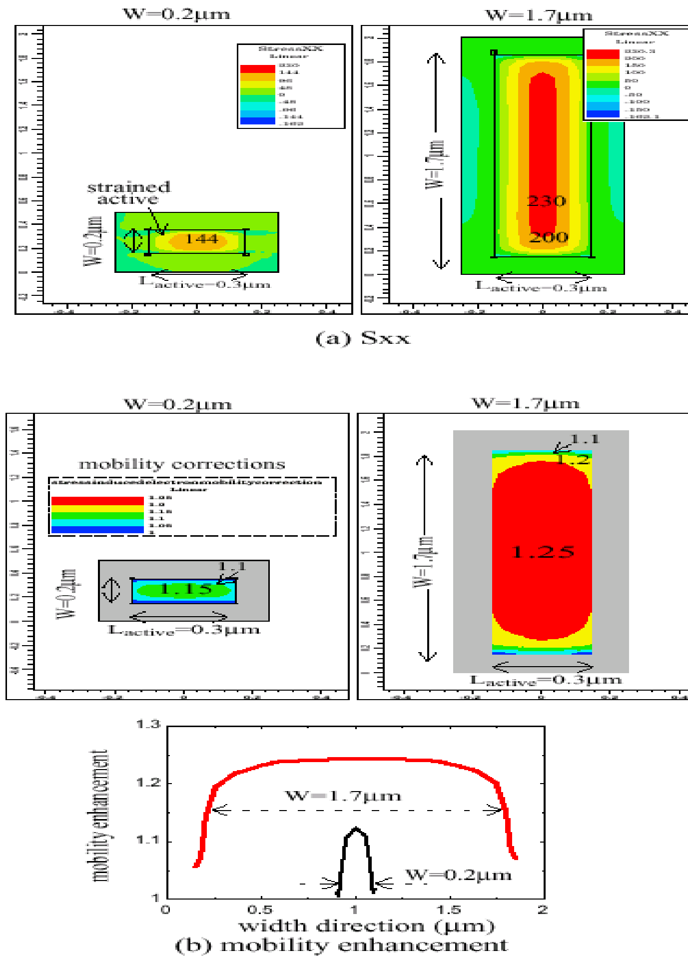



**FIGURE 7.33**

STI effects for narrow width active strained region, (a) $S_{xx}$ contours for W/L = 0.2 $\mu$m/0.1 $\mu$m and 1.7 $\mu$m/0.1 $\mu$m (b) $S_{zz}$ stress-induced mobility enhancements for W = 0.2 and 1.7 $\mu$m. After R. W. Dutton and C.-H. Choi, Presentation in NSTI Nanotechnology Conference 2004.

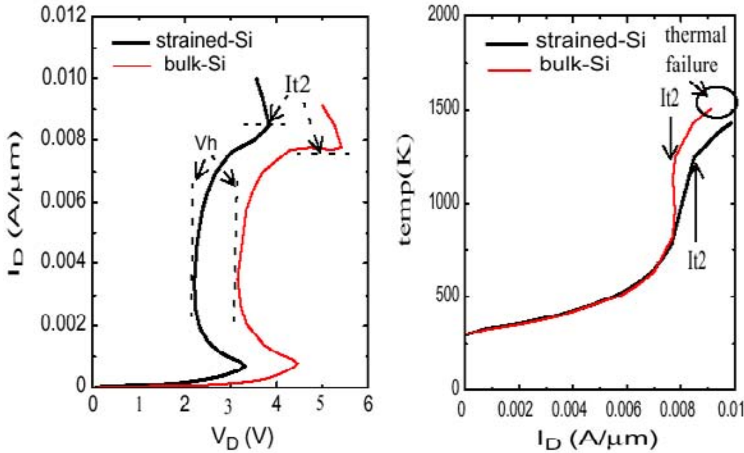



**FIGURE 7.34**
$I_d$-$V_d$ vs. temperature for strained-Si nMOS and bulk-Si nMOS. (a) $V_d$ vs. lattice temperature and (b) $I_d$ vs. lattice temperature. After R. W. Dutton and C.-H. Choi, Presentation in NSTI Nanotechnology Conference 2004.

### 7.5.1 Layout Dependence

With the downscaling of CMOS technology, MOSFET characteristics become more sensitive to the device layout pattern [39, 40]. In a conventional STI process, gap-fill oxide (TEOS) is deposited after dry etching of trench and the trench wall oxidation of typical 30 nm thickness at 1000$^o$C. During high-temperature densification at 1000$^o$C, oxides grow inevitably at the STI wall and increases its volume, and thus mechanical stress is generated at the shallow trench isolation edge [41]. This mechanical stress in the device affects the device characteristics in various ways [42, 43, 44, 45], for example, carrier mobility, hot carrier immunity, and dopant diffusion. STI-induced stress also causes crystal defects and increases the junction leakage [41]. There are some reports on enhancing device performance using mechanical stress [46, 47].

Due to the presence of stress in a MOSFET, the threshold voltage, transconductance, and saturation drain current are affected by the layout pattern change in spite of the same drawn gate length and width. This will result in a significant influence on the precise analog and logic circuits with many layout patterns, because the device characteristics are now different even though the circuit design parameters are the same and the device parameter cannot be described in a single parameter set. Thus the active-area layout dependence of CMOS device characteristics studies has become very important. The layout dependence has been investigated using mechanical stress simulation by Miyamoto et al. [48]. An approach to reduce shallow trench isolation-induced mechanical stress and its impact on reducing active-area layout dependence

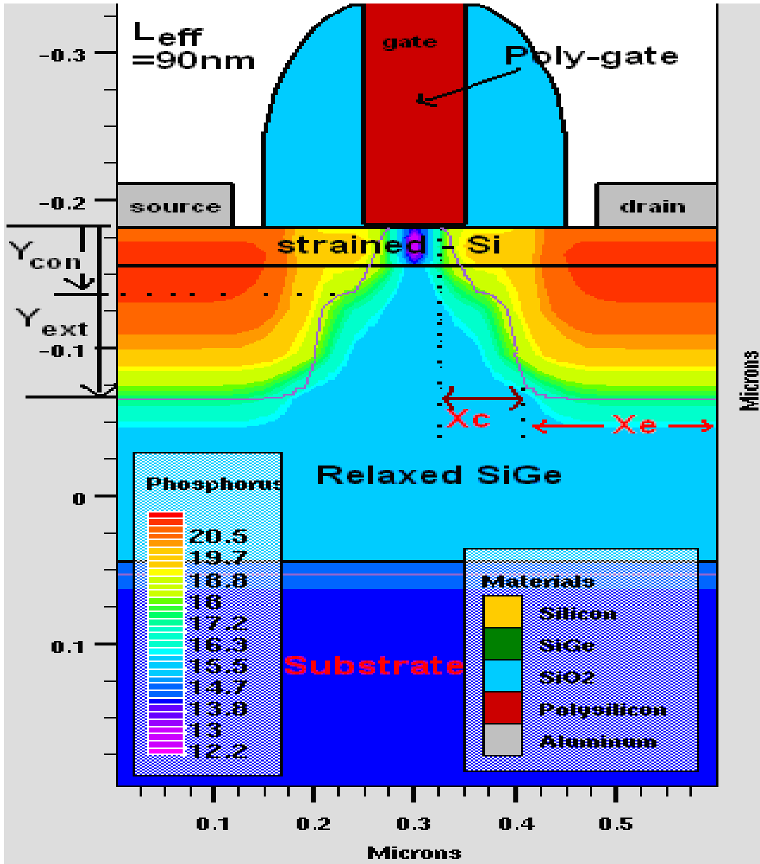



**FIGURE 7.35**
Simulated device structure obtained from ATHENA simulation for an n-channel strained-Si MOSFET. After A. R. Saha and C. K. Maiti, unpublished data.

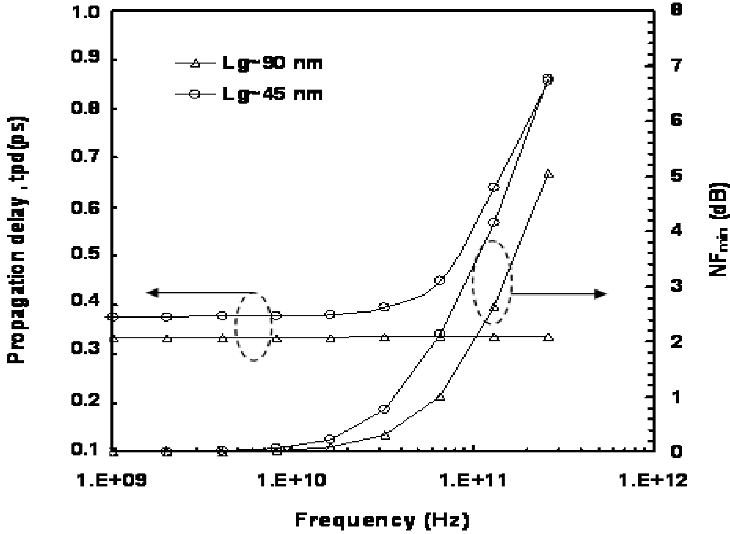



**FIGURE 7.36**
Simulated current gain vs. cutoff frequency as a function gate length for n-channel strained-Si MOSFETs. After A. R. Saha and C. K. Maiti, unpublished data.

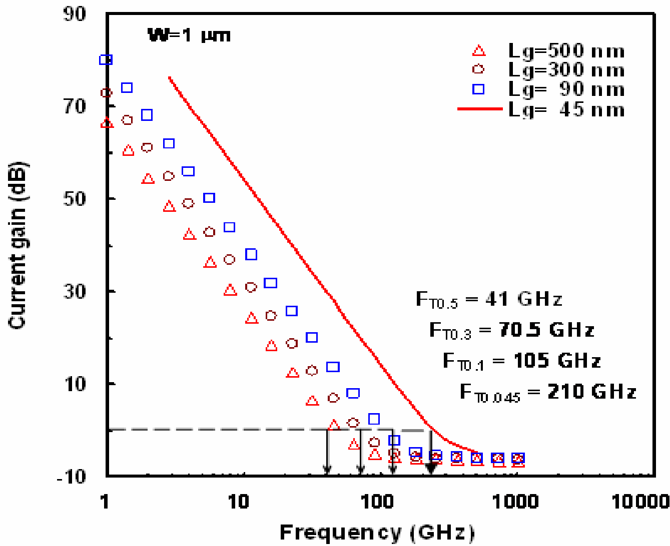



**FIGURE 7.37**
Simulated noise and delay performances for n-channel strained-Si MOSFETs. After A. R. Saha and C. K. Maiti, unpublished data.

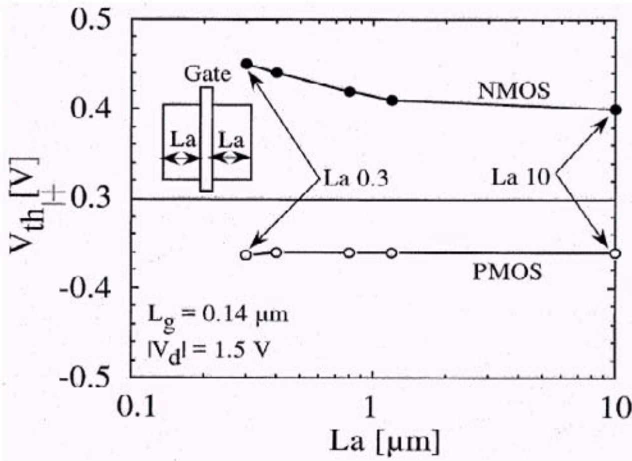



**FIGURE 7.38**
Dependence of threshold voltage on active-area length ($L_a$) in 0.18 $\mu$m CMOS technology. Here, L and W are constant ($L_g$ = 0.14 $\mu$m) and $W_g$ = 15 $\mu$m) with $L_a$ variation. After M. Miyamoto et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(440-443). With permission.

of MOSFET characteristics has also been reported.

In the layout parameter, active-area length (length from the gate edge to the STI edge, $L_a$) has significant influence on device characteristics. Also, the influence of layout pattern is different for n- and p-MOSFETs. Active-area length dependence of 0.18 $\mu$m CMOS technology ($L_g$ = 0.14 $\mu$m) is shown in Fig. 7.38. The $V_{th}$ of the n-MOSFET increases with decreasing $L_a$, in spite of the same gate length (0.14 $\mu$m) and width (15 $\mu$m). In contrast, the $V_{th}$ of the p-MOSFET is almost independent of $L_a$. The device characteristics $I_{ds}$-$V_{th}$ plot of nMOS and pMOS transistors with $L_a$ = 0.3 $\mu$m and $L_a$ = 10 $\mu$m are compared in Fig. 7.39. In n-MOSFET with $L_a$ = 0.3 $\mu$m, the $V_{th}$ are about 50 mV higher than that of the $L_a$ = 10 $\mu$m devices. However, the $I_{ds}$ of both devices appeared to be the same when the $V_{th}$ is adjusted because these plots are on the same $I_{ds}$-$V_{th}$ line. In contrast, the $V_{th}$ distribution of p-MOSFET is almost the same for both devices. The origin of increasing $V_{th}$ in n-MOSFET with smaller $L_a$ is that the large compressive stress originating in the STI edge reduces the diffusion of pocket ion implants (Boron) and the doping concentrations in channel edges become higher. The mechanism of these phenomena is schematically shown in Fig. 7.40.

Large compressive stress originating at the STI edge spreads to the center of the transistor region. The simulated $L_a$ dependence of longitudinal compressive stress in x-direction at the center of the channel region is shown in Fig. 7.41. The stress increases rapidly with decreasing $L_a$ in a critical length.

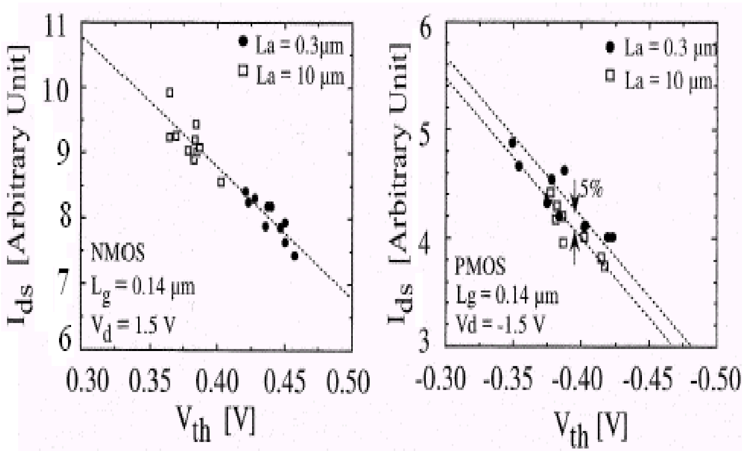



**FIGURE 7.39**
$I_{ds}$-$V_{th}$ plot of nMOS and pMOS transistors with $L_a = 0.3$ $\mu$m and $L_a = 10$ $\mu$m. After M. Miyamoto et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(440-443). With permission.

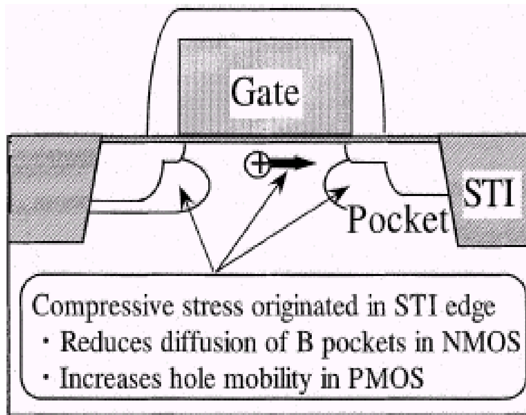



**FIGURE 7.40**
Schematic mechanism of increasing $V_{th}$ in n-MOSFET and increasing $I_{ds}$ in p-MOSFET with decreasing $L_a$. After M. Miyamoto et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(440-443). With permission.

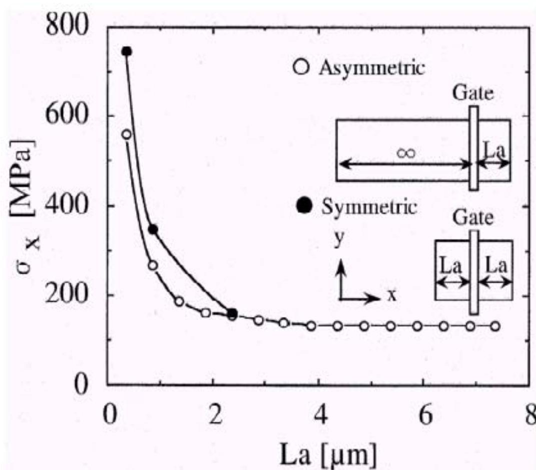



**FIGURE 7.41**
Simulated dependence of compressive stress in x direction on $L_a$ at the center of channel region. In case of asymmetric $L_a$ layout, only stress from one side STI edge is accounted. In case of symmetric layout, stress from both sides of STI edge is considered. After M. Miyamoto et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(440-443). With permission.

In the asymmetric layout as shown in the schematic in Fig. 7.41, only the stress from one STI edge is accounted. In the symmetric layout, the stresses from both sides of STI edges are added up and become larger than that in the asymmetric layout. The symmetric $L_a$-0.3 $\mu$m device has compressive stress of 750 MPa, which is about 600 MPa higher than that of the $L_a$-10 $\mu$m device.

### 7.5.2 Thickness Dependence

Biaxial tensile strained-Si is achieved from the growth of a thin, 100-200 Å, silicon layer on top of fully or partially relaxed-SiGe due to the lattice mismatch between the layers. Fiorenza et al. [49] have studied the effect of the strained-Si layer thickness on the characteristics of strained-Si MOSFETs on SiGe virtual substrates. n-MOSFETs were fabricated on strained-Si substrates with various strained-Si thicknesses, both above and below the strained-Si critical thickness. The low-field electron mobility and subthreshold characteristics of the devices were measured. It has been shown that the low-field electron mobility is increased by about 1.8 times on all wafers and is not significantly degraded on any of the samples, even for a strained-Si thickness far greater than the critical thickness. From the subthreshold characteristics, it was observed that the off-state leakage current is greatly increased for the devices with a strained-Si thickness that exceeds the critical layer thickness.

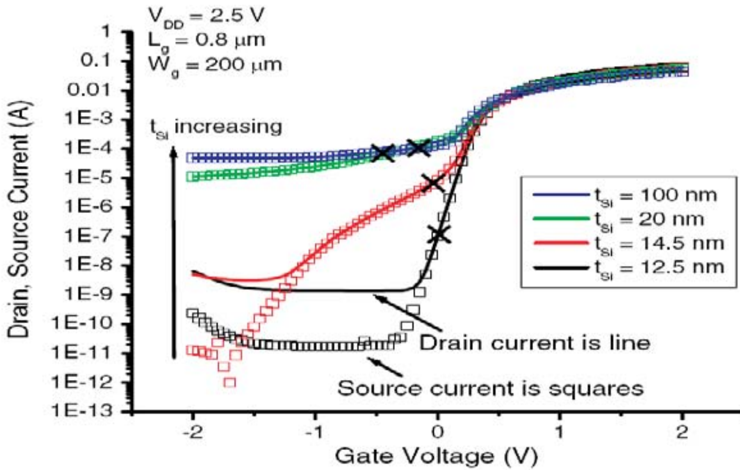



**FIGURE 7.42**

Subthreshold characteristics of strained-Si MOSFETs with a strained-Si thickness above and below the critical thickness. The devices that are above critical thickness (14.5 nm, 20 nm, and 100 nm) show greatly increased levels of off-state leakage current, while the device below the critical thickness (12.5 nm) does not. Both the source current and the drain current are shown and they are equal in the elevated leakage regime. The leakage path is known to extend from the drain to the source, and not from the drain to the substrate. After J. G. Fiorenza et al., *Semicond. Sci. Technol.*, Vol. 19, 2004(L4-L8).

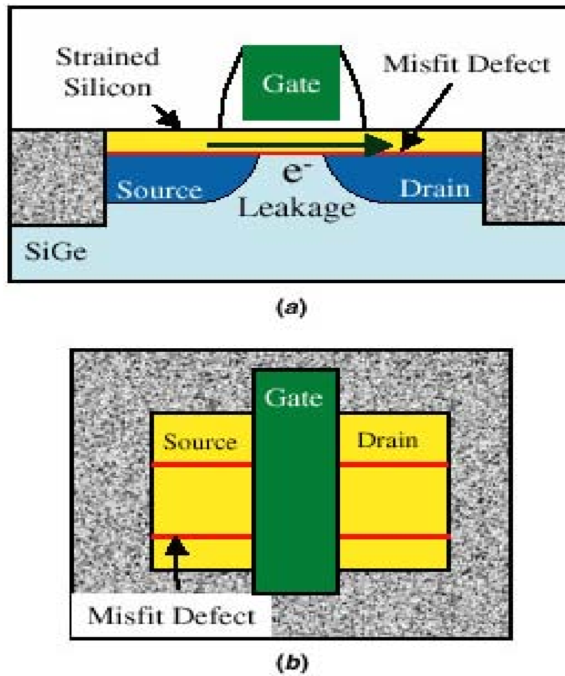



**FIGURE 7.43**
(a) Schematic views of a strained-Si MOSFET, showing the drain current leakage mechanism in strained silicon films containing misfit dislocations. (b) Misfit dislocations at the Strained-Si/SiGe interface form dislocation pipes and result in highly localized points of current. After J. G. Fiorenza et al., *Semicond. Sci. Technol.*, Vol. 19, 2004(L4-L8).

The subthreshold characteristics of transistors ($L_g = 0.8\ \mu$m, W = 200 $\mu$m) on each of the four Strained-Si wafers are compared in Fig. 7.42. The devices on the wafer with 12.5 nm of strained-Si show normal behavior: the drain current falls exponentially with gate voltage for a gate voltage moderately below threshold and then settles to a constant current floor for a gate voltage far below threshold. However, the devices fabricated on wafers with strained-Si thicker than $t_{crit}$ show aberrant behavior. They suffer from greatly elevated levels of off-current leakage, and the magnitude of the leakage current increases with the thickness of the strained silicon beyond the $t_{crit}$.

The authors have also explained the mechanism of the elevated leakage current observed in the Strained-Si MOSFETs beyond the critical thickness by considering the position of the misfit dislocations in the structure of the MOSFET, as shown in Fig. 7.43. The misfit dislocations form at the silicon SiGe interface and can cross between a MOSFET source and drain. These misfit dislocations form dopant diffusion pipes that cause localized source-to-drain shorts, resulting in increased off-current levels. There are several alternative explanations for the increased off-state leakage current. Misfit dislocations could cause leakage between the drain and the substrate, not between the drain and the source or the leakage could occur due to direct conduction along a dislocation, and not involve dopant diffusion in any way.

Most strained-Si MOSFET results published to date utilize substrates with strained-Si over the critical thickness. It is very important to establish the correct limit on the Strained-Si thickness, and to understand the implications on MOSFET performance of the use of supercritical films. In general, the strained-Si thickness should be below the strained-Si critical thickness to avoid strained-Si relaxation and mobility degradation.

### 7.5.3 Orientation Dependence

As the hole mobility on (110) surface is about two times as high as that on (100) surface [50, 51, 52, 53], p-MOSFETs on a (110) surface will be useful for the improvement of drain current drive and for realizing high-speed CMOS circuits, as well as reducing the imbalance between n- and p-channel drive current. Although the hole mobility on (110) surface is about two times as high as that on (100) surface, the electron mobility of (110) n-MOSFETs is much lower and about one half of that of (100) counterpart [50, 53, 54]. However, the lower electron mobility is a serious problem for the enhancement of CMOS current drive. As a result, it is important to develop a new device technology with enhanced electron mobility in the (110)-surface CMOS.

High-speed strained-Si silicon-on-insulator CMOS devices on (100) surface, using the strained-Si channel with high carrier mobility on relaxed SiGe-on-insulator (SGOI) substrates, have been demonstrated by several research groups [37]. Strained-SOI n-MOSFETs, successfully fabricated on (110)-surface, show an electron mobility enhancement of 23% along the (001) direction, against that of the (110)-surface unstrained-bulk MOSFETs [54]. As a

result, the electron mobility ratio of (110) strained-SOI n-MOSFETs to the universal mobility of (100) bulk MOSFETs increases up to 81%.

Mizuno et al. [55] have reported n-MOSFETs on (110)-surface strained-SOI. The (110) strained-SOI substrates can be realized by epitaxially growing 30-nm-thick strained-Si films on (110)-surface relaxed SGOI structures. A 200-nm thick strained-SiGe layer with the Ge content of 10% was epitaxially grown on the (110) SOI wafers, where the SOI and the buried-oxide thicknesses were 35 and 300 nm, respectively. The strained-Si channel was formed on (110) relaxed-SGOI substrates, fabricated by the germanium condensation technique to (110)-surface SOI substrate with the (111) cut plane. The 130-nm thick (110) relaxed-SGOI substrates were realized by oxidizing this strained-SiGe layer of the Ge condensation technique of SiGe layer at 1200$^o$C [56]. The Ge content of the (110) SGOI substrate was 25%, and the relaxation rate was about 64%, resulting in the effective Ge content of 16%, which is equivalent to the lateral lattice constant of fully relaxed-SGOI substrates with x of 16%.

The subthreshold swing of (110) strained-MOSFETs were 160 mV/dec, although the subthreshold swing of (110) bulk devices is about 70 mV/dec. From the subthreshold swing values of strained-MOSFETs, the gate-oxide strained-Si interface state density is estimated to be about $2\times10^{12}$ cm$^{-2}$ eV$^{-1}$, which is due to the quality of the (110) strained-SOI substrate without any optimization. The drain leakage current of (110) strained-SOIs was lower than 0.1 pA at the drain bias of 1.5 V.

Fig. 7.44 shows the behaviors of (110) strained-SOI MOSFETs (closed characters) and (110) unstrained-bulk-MOSFETs (open characters) as a function of the effective electric field. As shown as the inset, the current flow direction of (110) strained-SOI and (110) bulk-MOSFETS are 10 and 0$^o$ tiled from the <001> direction, respectively. The solid line shows the electron universal mobility of (100) bulk MOSFETs [57]. The electron mobility was measured from the drain currents at $V_d$ = 5 mV, and the gate-to-channel capacitance in various current flow directions, using n-MOSFETs with $W_{eff}/L_{eff}$ = 102 $\mu$m/200 $\mu$m. The mobility enhancement in (110) strained-SOI MOSFETs is higher in the whole range of $E_{eff}$ than that of (110) bulk-MOSFETs. The maximum mobility enhancement factor against the $\mu_{eff}$ of the (110) bulk-MOSFETs is 23% at $E_{eff}$ of around 0.3 MV/cm. The slope of the $\mu_{eff}$-$E_{eff}$ characteristics of (110) strained-SOI devices, which is almost the same as that of (110) bulk-MOSFETs and is larger than that of (100) devices, also indicates the larger contribution of surface roughness scattering at (110) surface orientation [54]. The electron mobility enhancement strongly depends on the current flow direction (see Fig. 7.45) and has the maximum value of 23% along the <001> direction, which can be explained by the anisotropic conductivity mass of the four-fold valleys. The electron mobility of strained-SOI MOSFETs has the maximum value along the <001> direction, and the mobility enhancement against that of the (110) bulk MOSFETs amounts to 23%.

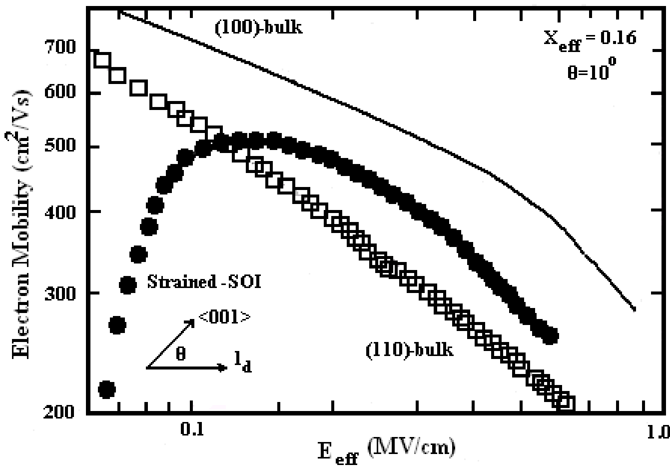



**FIGURE 7.44**

Effective electron mobility as a function of (closed circles) transverse effective field of (110) strained-SOI and (open squares) (110) unstrained-bulk MOSFETs for inversion-layer in nMOS, where the effective Ge content is 16%. After T. Mizuno et al., *IEEE Electron Dev. Lett.*, Vol. 24, 2003(266-268). With permission.

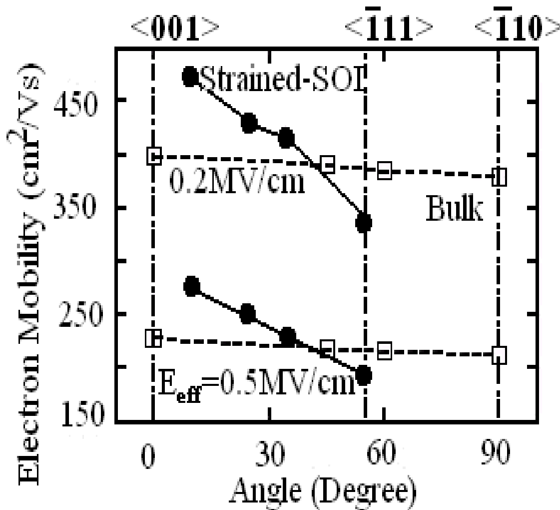



**FIGURE 7.45**

(110) electron mobility (closed circles) of strained- and bulk-MOSFETs (open squares) vs. the current flow direction tiled from the <001> direction, where $E_{eff}$ is 0.2 and 0.5 MV/cm. The upper x-axis shows the equivalent direction of the current flow direction. After T. Mizuno et al., *IEEE Electron Dev. Lett.*, Vol. 24, 2003(266-268). With permission.

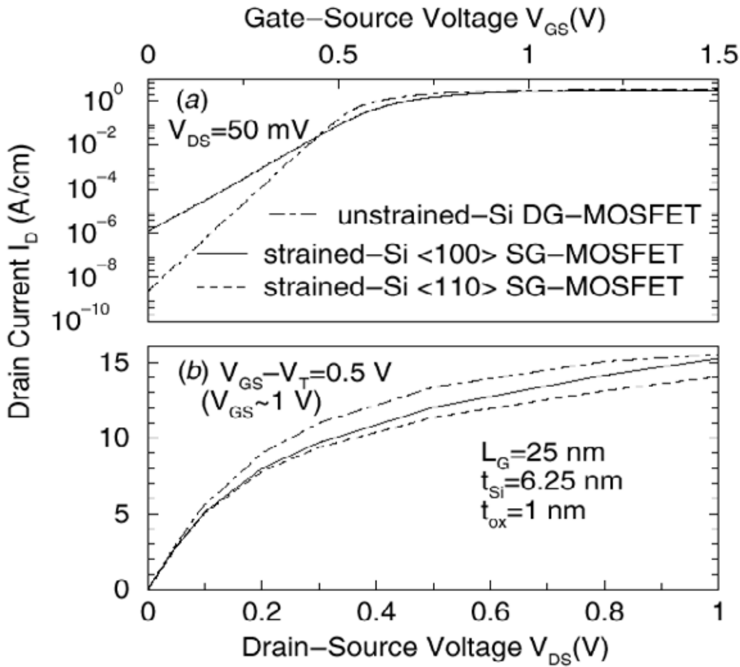



**FIGURE 7.46**
Transfer characteristics according to the DD model (a) and output characteristics computed by Monte Carlo simulation (b) of an unstrained-Si DG-MOSFET and a strained-Si SG-MOSFET with a channel oriented along the crystallographic in-plane (110) or (100) direction. Note that there is no difference between the (110) and the (100) directions in the strained-Si transfer characteristics at $V_{ds}$ = 50 mV (a), because the anisotropy only appears at higher drain voltages (see (b)). After F. M. Bufler et al., *Semicond. Sci. Technol.*, Vol. 19, 2004(S122-S124).

### 7.5.4 Hetero-FETs: Single Gate vs. Double Gate

Bufler et al. [58] have used self-consistent full-band Monte Carlo simulations to compare the performance of nanoscale strained-Si single-gate (SG) and unstrained-Si double-gate (DG) MOSFETs for a gate length of 25 nm. It has been show that almost the same on-current as in the DG-MOSFET can be achieved by strain in a SG-MOSFET for the same gate overdrive. This is due to the compensation of the higher electron sheet density in the two inversion channels of the DG-MOSFET by the higher strain-enhanced velocity in the channel of the SG-MOSFET. The on-current of the strained-Si SG-MOSFET is almost 10% larger for a channel orientation along the crystallographic (100) direction than for the (110) direction.

Fig. 7.46(a) shows the transfer characteristics where the Ge content and

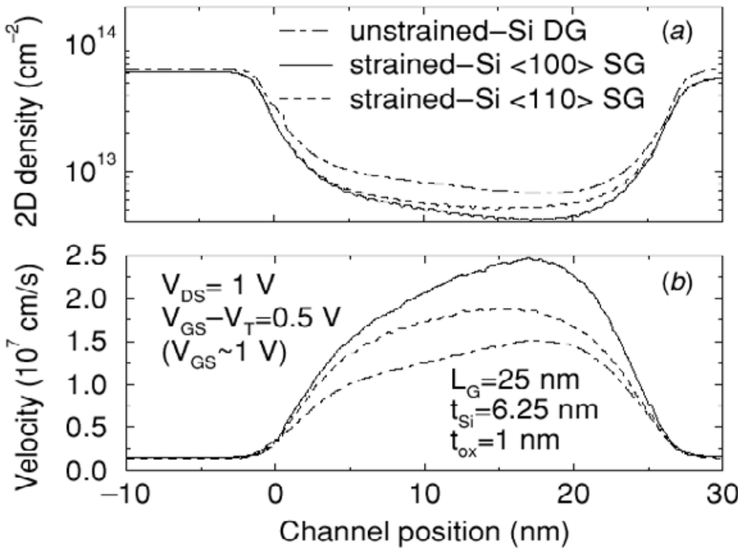



**FIGURE 7.47**
Profiles along the channel of (a) the sheet density obtained by integration of the electron density perpendicular to the $Si/SiO_2$ interface over the silicon film thickness $t_{Si}$ and (b) the drift velocity, averaged perpendicularly to the interface with the electron density, in the on-state of the three device configurations of Fig. 7.46. After F. M. Bufler et al., *Semicond. Sci. Technol.*, Vol. 19, 2004(S122-S124).

the doping level in the SiGe buffer were adjusted to yield, similar to the DG MOSFET, a threshold voltage of about $V_t \sim 0.5$ V. The off-current in the strained-Si MOSFET is about three orders of magnitude larger than in the DG MOSFET, but the value of 0.1 nA/$\mu$m is still very small and appropriate for high-speed applications. The output characteristics in Fig. 7.46(b), computed for the same gate overdrive of $V_{gs}$-$V_t = 0.5$ V, demonstrate that almost the same on-current $I_{on}$ as in the DG MOSFET can be achieved by strain in a SG MOSFET.

Fig. 7.47 shows the profiles of sheet density and drift velocity It is observed that the higher charge density of the two inversion channels in the DG MOSFET is compensated by a higher drift velocity in strained-Si. This confirms that the on-current is determined by quasiballistic transport, because the maximum enhancement of the in-plane velocity in bulk (001)-strained-Si in the 100 direction is 5% at medium electric fields (the low-field mobilities and saturation velocities are the same and the difference in the thermal injection velocities is negligibly small), whereas the transient bulk velocity overshoot peak is 30% larger in the (100) direction. This simulation study shows a high

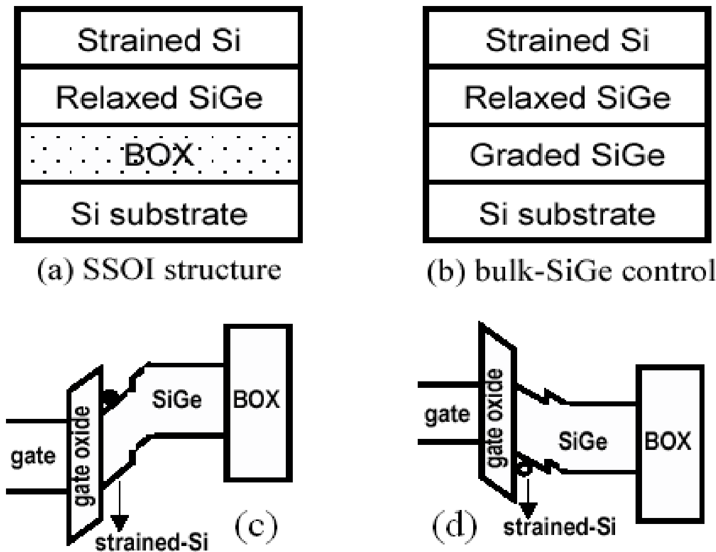



**FIGURE 7.48**
Schematic layer sequence and band diagram for SSOI: (a) SSOI structure: strained-Si on SGOI, (b) bulk-SiGe control: strained-Si/relaxed-SiGe on bulk-Si substrate, (c) band diagram of SSOI under n-FET bias condition, and (d) band diagram of SSOI under p-FET bias condition, where holes are populated on the surface strained-Si layer (surface channel). After Z. Cheng et al., SMA Symposium 2003.

potential of both DG and strained-Si MOSFETs for the nanoscale regime. However, to decide which of the two approaches is to be preferred will require more detailed investigations considering also, e.g., restrictions on $I_{off}$ or fabrication-related issues.

### 7.5.5 Hetero-FET: Dual Channel Structure

Various heterostructure substrates have been used to investigate biaxial strain and high-mobility channel materials, including epitaxial structures on bulk substrates and silicon-on-insulator implementations. Relaxed-SGOI is a versatile substrate that can be used to integrate various device structures, such as strained-Si devices on SGOI, pure Ge devices on SGOI, and III-V optoelectronics on SGOI, utilizing SGOI's capability of matching the lattice constant of SiGe with other materials. Surface-channel SSOI structure has been considered as a promising structure for high-mobility CMOS [59, 60, 61, 62], with enhancement of both electron and hole mobilities. Several workers [63, 64] have used tensile strained-Si layer grown on relaxed-SGOI substrate (SSOI

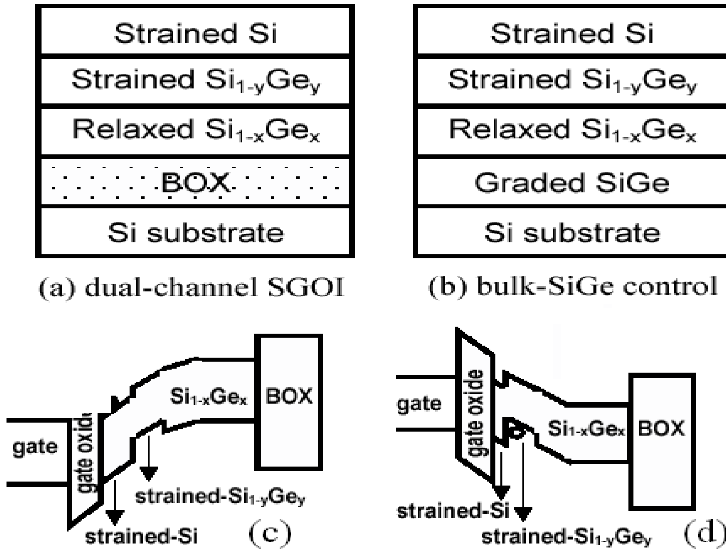



**FIGURE 7.49**
Schematic layer sequence and band diagram for dual-chanel SGOI: (a) dual-channel SGOI structure: strained-Si and strained-$Si_{1-y}Ge_y$ on SGOI, (b) bulk-SiGe control: strained-Si/ strained-$Si_{1-y}Ge_y$/relaxed-$Si_{1-x}Ge_x$ (y > x) on Si substrate, (c) band diagram of dual-channel SGOI under n-FET bias condition where an electron channel forms on the surface strained-Si layer (surface channel), and (d) band diagram of dual-channel SGOI under p-FET bias condition where the hole channel forms on the buried strained-$Si_{1-y}Ge_y$ layer (buried channel). After Z. Cheng et al., SMA Symposium 2003.

structure) to obtain significant electron mobility enhancement in surface channel strained-Si n-MOSFETs. To further boost hole mobility, compressively strained-$Si_{1-y}Ge_y$ grown on relaxed-$Si_{1-x}Ge_x$ structure (y > x) can be used for p-MOSFETs [65]. An excellent review the status of research on Si/SiGe heterostructure MOSFETs in bulk and on-insulator technologies involving substrate strain may be found in reference [66].

A dual-channel MOSFET structure combining a tensile strained-Si layer for n-MOSFETs and a compressively strained-$Si_{1-y}Ge_y$ layer for p-MOSFETs has been proposed for CMOS [67]. Such a dual-channel structure on bulk relaxed-SiGe substrate has shown a significant hole mobility enhancement (up to 5.15 times at field of 0.3 MV/cm for 80% strained-SiGe on 50% relaxed-SiGe) in p-MOSFETs. In the following, we discuss the performance of both n- and p-MOSFETs fabricated on these two structures. The n-MOSFETs require a thick enough strained-Si layer for its electron surface channel. The p-MOSFETs, on the other hand, require the surface strained-Si layer to be

very thin, otherwise parasitic hole channel will form in this strained-Si layer at high gate bias.

In a SSOI structure, a tensile strained-Si layer is grown on a relaxed SGOI substrate as shown schematically in Fig. 7.48(a) where the relaxed-SiGe layer has Ge fraction of 22%. The surface-channel CMOS devices built on SSOI structure provide enhancement on both electron and hole mobilities. The corresponding band diagrams of SSOI under n-FET and p-FET bias condition are illustrated in Figs. 7.48(c) and (d), respectively. Electrons and holes are both populated at the surface strained-Si layer for n- and p-MOSFETs, respectively. As surface-channel devices, SSOI CMOS has advantage on the subthreshold slope.

Unlike surface-channel SSOI CMOS discussed above, dual-channel SGOI structures utilize different layers for electron channel and hole channel. As shown in Fig. 7.49(a), the surface strained-Si layer is used for n-MOSFETs (surface channel), and the underneath strained-$Si_{1-y}Ge_y$ layer for p-MOSFETs (buried channel). In this structure, a compressively strained $Si_{0.4}Ge_{0.6}$ layer and then a tensile strained-Si layer are grown on relaxed-SGOI substrate where the relaxed-SiGe layer has a Ge fraction of 30%. The corresponding band diagrams under n-FET and p-FET bias condition are illustrated in Figs. 7.49(c) and (d), respectively. This structure features high hole mobility enhancement. Starting with a relaxed-$Si_{0.7}Ge_{0.3}$ buffer, for example, a strained-Si layer provides n-MOSFETs over 80% mobility enhancement, while a strained-$Si_{0.4}Ge_{0.6}$ (60% Ge fraction) layer gives p-MOSFETs about 200% hole mobility enhancement as demonstrated by experiment in bulk-SiGe substrate [67].

From device performance point of view, the surface-channel SSOI gives good subthreshold characteristics. The dual-channel SGOI structure, on the other hand, provides much higher hole mobility than SSOI structure. It also has the advantage to optimize strained-Si layer and strained-$Si_{1-y}Ge_y$ layer for nMOS and p-MOS separately. But the presence of the strained-$Si_{1-y}Ge_y$ layer may affect the performance of nMOS devices. From process point of view, selective etching the surface strained-Si layer on the p-MOSFET regions only and stopping at a thin Si thickness is a process challenge, and the process for dual-channel CMOS thus becomes more complicated than SSOI. Moreover, higher Ge fraction epi layer may subject to higher defect density during epi growth.

Fig. 7.50 shows the effective electron mobilities vs. effective field from n-MOSFETs on SSOI structure and the bulk-SiGe control structure (strained-Si/relaxed-SiGe) respectively, plotted with the universal mobility from reference [57]. The mobilities are extracted from drain current measurements with low drain voltage and with the inversion charge extracted from the gate-to-channel capacitance $C_{GC}$ from the same device. Typical drain current output characteristics of the n-MOSFETs on SSOI and on bulk-SiGe control structure are shown in Fig. 7.51. The subthreshold characteristics are shown in Fig. 7.52. The subthreshold swing are 84 mV/dec and 86 mV/dec for SSOI and

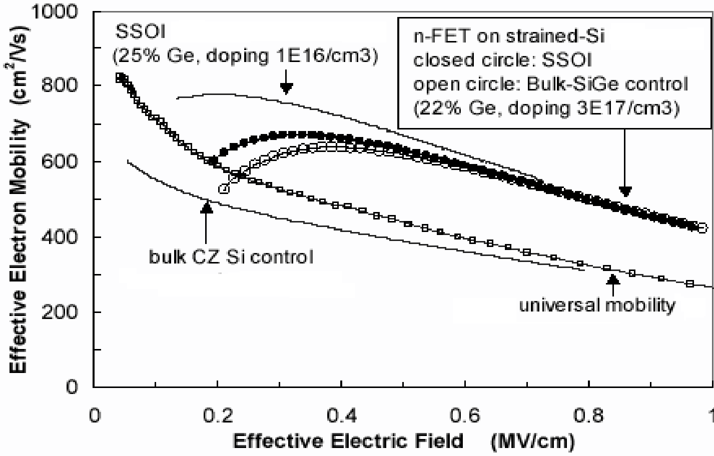



**FIGURE 7.50**
Measured effective electron mobility vs. effective electric field from n-MOSFETs on SSOI structure and bulk-SiGe control structure. The mobilities are extracted from drain current with the inversion charge extracted from the gate-to-channel capacitance. After Z. Cheng et al., SMA Symposium 2003.

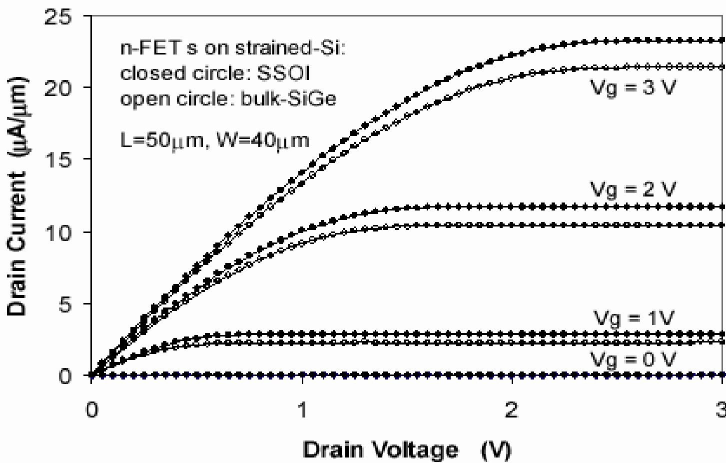



**FIGURE 7.51**
Drain current output characteristics: drain current vs. drain voltage for SSOI (closed circle), and bulk-SiGe control (open circle). Channel length: 50 $\mu$m, channel width: 40 $\mu$m. The SiGe body contact terminal is grounded during measurement. After Z. Cheng et al., SMA Symposium 2003.

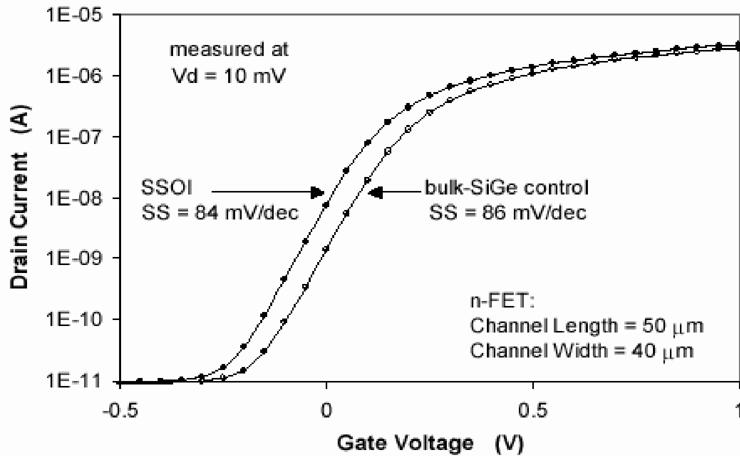



**FIGURE 7.52**
Subthreshold characteristics: drain current vs. gate voltage at low drain voltage for SSOI (open circle), and bulk-SiGe control (closed circle). Channel length: 50 $\mu$m, channel width: 40 $\mu$m. The SiGe body contact terminal is grounded during measurement. After Z. Cheng et al., SMA Symposium 2003.

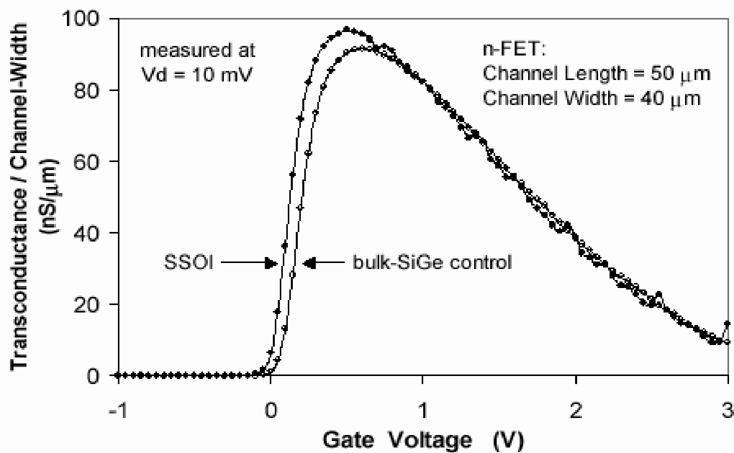



**FIGURE 7.53**
Transconductance characteristics divided by channel width. SSOI: open circle; bulk-SiGe control: closed circle. Channel length: 50 $\mu$m, channel width: 40 $\mu$m. The SiGe body contact terminal is grounded during measurement. After Z. Cheng et al., SMA Symposium 2003.

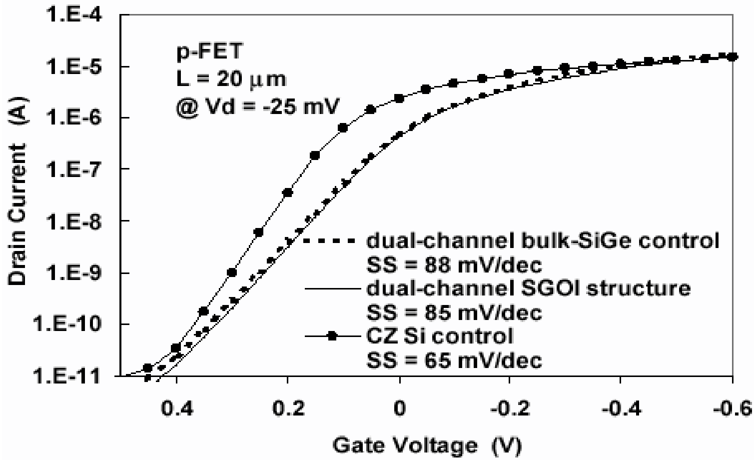



**FIGURE 7.54**
Subthreshold characteristics of dual-channel SGOI: drain current vs. gate voltage at low drain voltage for p-MOSFETs on dual-channel SGOI, dual-channel bulk-SiGe control, and CZ Si control. Channel length: 20 $\mu$m. The SiGe body contact terminal is grounded during measurement. After Z. Cheng et al., SMA Symposium 2003.

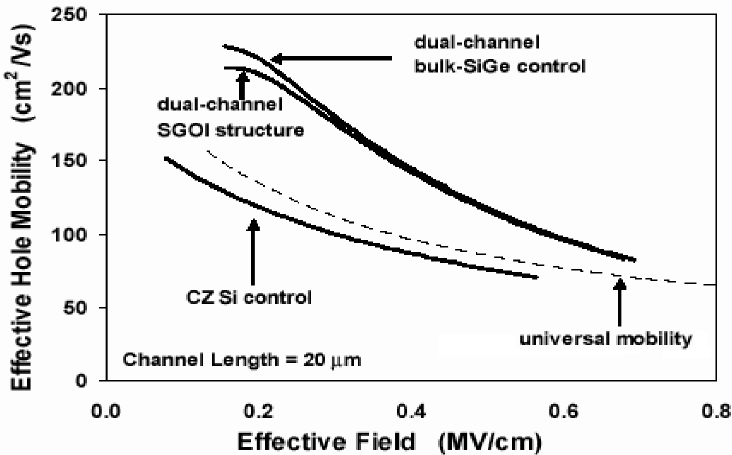



**FIGURE 7.55**
Measured effective hole mobility vs. effective field from p-MOSFETs on dual-channel SGOI, dual-channel bulk-SiGe control, and CZ Si control. The mobilities are extracted from drain current with the inversion charge extracted from the gate-to-channel capacitance. After Z. Cheng et al., SMA Symposium 2003.

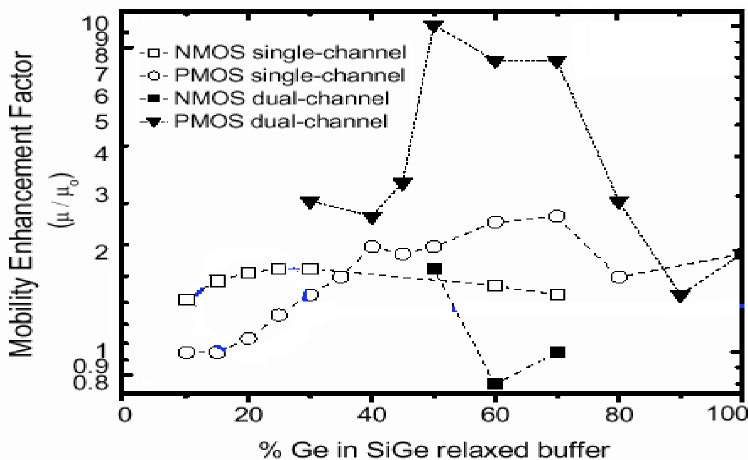



**FIGURE 7.56**
The mobility in enhancements in strained-SiGe n- and p-MOSFETs vs. relaxed buffer Ge concentration. Data for both single channel and dual channel structures are shown. Largest performance gain can be attained with SiGe dual-channel grown on $Si_{0.5}Ge_{0.5}$. Source: J. L. Hoyt, IFST 2004 Lecture.

bulk-SiGe control, respectively. The corresponding transconductance characteristics divided by channel width ($g_m$/W) are shown in Fig. 7.53. Fig. 7.54 shows the subthreshold characteristic of p-MOSFETs on dual-channel SGOI and that on dual-channel bulk-SiGe control structure, together with that on bulk CZ Si control. As a result of buried-channel device, the subthreshold swing, 85 mV/dec and 88 mV/dec for dual-channel SGOI and bulk-SiGe control, respectively, are larger than the bulk CZ Si control (65 mV/dec).

Fig. 7.55 shows the effective hole mobilities vs. effective field from p-MOSFETs on the dual-channel SGOI structure and on the bulk-SiGe control structure. The mobility from bulk CZ Si control, and the universal mobility from reference [57], are also shown in the figure for comparison. The plots show the performance of p-MOSFETs on dual-channel SGOI structure are very similar to that on the dual-channel bulk-SiGe control. The hole mobility enhancement over the bulk-Si control is about 90% at effective field of 0.2 MV/cm, and about 60% at 0.5 MV/cm. The enhancement reduces with the electric field increases, as also observed in the experiment on bulk-SiGe [67].

Although there is enhancement on hole mobility, this enhancement factor is smaller than what is expected based on the experiment on bulk-SiGe substrate. This is most probably due to the poor control of MOSFET fabrication process and thermal budget. The anneal step of 850$^o$C for 30 min is believed too high and may cause the strained-$Si_{1-y}Ge_y$ layer to lose its strain, leading to small enhancement on hole mobility. Largest performance gain

can be attained with SiGe dual-channel grown on $Si_{0.5}Ge_{0.5}$ (see Fig. 7.56). It may be noted that the highest enhancement observed in a dual channel, strained-Si/strained-Ge/relaxed-$Si_{0.5}Ge_{0.5}$ structure, demonstrating an electron mobility enhancement of 1.7 times for the nMOS, and about a 10× hole mobility enhancement for the pMOS. These data suggest that mobility can be enhanced by approximately 1000% over conventional Si MOS mobility, assuming that short-channel device optimization and low temperature processing is possible.

Table 7.2 shows the comparison between the two SGOI structures for CMOS applications: (i) the surface-channel SSOI structure and (ii) dual-channel SGOI structure. The comparison also holds for their two counterpart structures on bulk-SiGe substrates: i.e., the surface-channel strained-Si/relaxed-SiGe on Si substrate and the dual-channel strained-Si/strained-$Si_{1-y}Ge_y$/ relaxed-$Si_{1-x}Ge_x$ on Si substrate.

## 7.6 Strained-Si MOSFETs: Reliability

The study of the ionization phenomena is of great importance both from the practical and theoretical points of view. In fact, the current-voltage operating range and the noise figure of SiGe-based devices are adversely affected by the onset of impact ionization multiplication, which is enhanced due to the low energy gap of this material.

Strained-Si technology holds great promise in the continuing drive to push CMOS along the scaling roadmap [68]. Currently, research in this area is mainly focused on logic applications. In mixed-signal applications, such as for wireless communication applications, where a high-frequency power amplifier function is essential, impact ionization becomes very important because it ultimately determines the breakdown voltage and subsequently determines the maximum power that the technology can deliver. At this time, little is known about the potential of strained-Si technology for mixed-signal applications. Impact ionization is also a significant consideration in device design for most other applications as it greatly affects device reliability.

Furthermore, as opposed to the behavior observed in most semiconductors, a reduction of the common-emitter breakdown voltage $BV_{CEO}$ with increasing temperature has been observed in SiGe HBTs [69]. The understanding of this behavior is of great importance due to the possible onset of a positive feedback loop between power dissipation and ionization-induced increase of output current. This may also be an important issue in electro-thermal modeling of SiGe-based devices.

Avalanche multiplication characteristics of $Si_{1-x}Ge_x$ photodiodes for the entire composition range $0.08 < x < 1$. The impact ionization coefficients have

**TABLE 7.2**
Comparison of Surface-channel SSOI and dual-channel SGOI structures. After Z. Cheng et al., SMA Symposium 2003.

| Structure | Surface-channel SSOI structure: strained-Si on SGOI | Dual-channel SGOI structure: strained-Si/ strained-$Si_{1-y}Ge_y$ on SGOI |
|---|---|---|
| Electron mobility | High | High |
| Hole mobility | Moderate | High |
| Channel: surface or buried | Surface channel for both n- and p-MOSFETs | Surface channel for n-MOSFETs Buried channel for p-MOSFETs |
| Strained-Si thickness | Same for n-MOSFETs and p-MOSFETs | Thicker for n-MOSFETs, thinner for p-MOSFETs |
| Parasitic channel for p-MOSFETs | Parasitic channel may form in the relaxed-SiGe layer at low bias | Parasitic channel may form in the surface strained-Si layer at high bias: The strained-Si layer needs to be very thin to avoid parasitic channel |
| Process complexity | Simpler | More steps: thinning Si layer in pMOS region |
| Optimization for nMOS and pMOS | Can't optimize separately | Can optimize separately: strained-Si for nMOS, and strained-$Si_{1-y}Ge_y$ for p-MOS |

been calculated from measured multiplication coefficients. Theoretical Monte Carlo simulations of impact ionization in strained-Si have also been carried out. Both studies show an increase in the impact ionization coefficient over that of bulk-Si. Experimental results of impact ionization coefficients in Si and Ge show that Shockley's model [70] is applicable to these semiconductors.

The impact ionization process depends largely on the band structure and bandgap. Impact ionization occurs when highly energetic charge carriers excite an electron to the conduction band (or hole to the valence band) to create an electron-hole pair. The electron-hole pair generation rate is given by [71]

$$G = \alpha_n n v_n + \alpha_p p v_p \tag{7.8}$$

where $\alpha_n$ is the electron ionization rate defined as the number of electron-hole pairs generated by an electron per unit distance, $\alpha_p$ is the analogous hole ionization rate, n is the electron concentration, p is the hole concentration, and $v_n$ and $v_p$ are the electron and hole saturation velocities, respectively. A simplified, local field-based empirical expression for the ionization rate is given by [72]

$$\alpha = Aexp(-b/E) \tag{7.9}$$

where A and b are experimentally determined constants and E is the electric field. Eqn. 7.9 shows that the carrier ionization rate is exponentially dependent on the electric field. Impact ionization in SiGe has been experimentally studied by means of photodiodes [73].

Waldron et al. [74] have experimentally studied the impact ionization in a strained-Si/SiGe heterostructure. It has been found that, unlike bulk-Si, impact ionization in this heterostructure has a positive temperature coefficient. Coupled with the severe self-heating that characterizes this heterostructure, this results in significantly higher levels of impact ionization when compared with bulk-Si devices operating under identical conditions. The authors contend that in strained-Si, bandgap narrowing, reduced effective mass, and a longer mean free path lead to a positive temperature coefficient for impact ionization.

Fig. 7.58 shows the temperature dependence of the source and body currents and impact ionization for both strained-Si/SiGe and bulk-Si. Impact ionization in the strained-Si/SiGe sample exhibits a strong positive dependence on temperature, which is opposite to what is seen in bulk-Si. The positive temperature dependence of the impact ionization is also apparent in the breakdown performance of the devices. This is the result of the interplay between the severe self-heating that occurs as a result of the low thermal conductivity of the SiGe buffer coupled with a strong positive temperature dependence of impact ionization in this material system. At 200°C the breakdown voltage of the strained-Si/SiGe devices decreases by 2 V; whereas, it increases by 0.3 V for the bulk-Si.

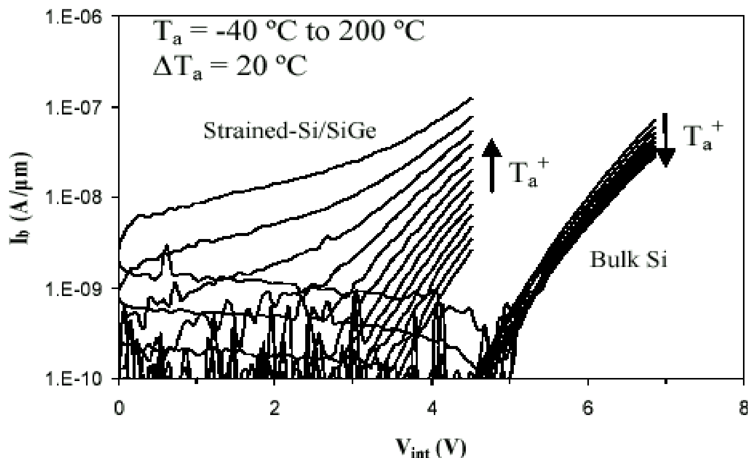



**FIGURE 7.57**
$I_b$ vs. $V_{int}$ for the range of ambient temperatures between -40$^o$C and 200$^o$C. The strained-Si/SiGe sample exhibits a strong positive temperature dependence in contrast to bulk-Si, which has a much weaker negative dependence. Leakage current in both structures increases at higher temperatures. After N. S. Waldron et al., *IEEE IEDM Tech. Dig.*, 2003(813-816). With permission.

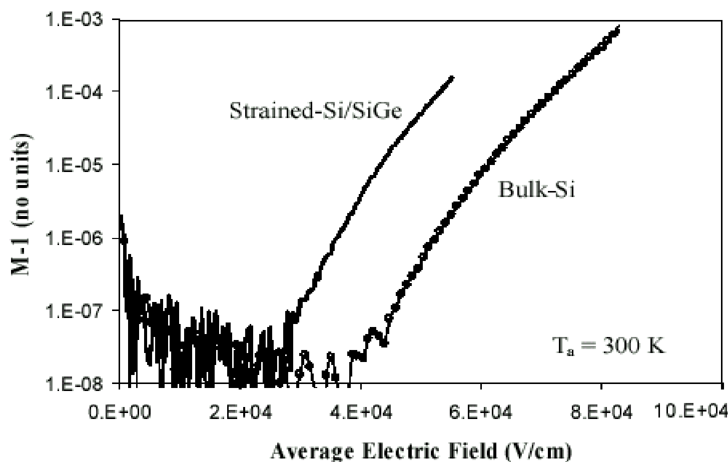



**FIGURE 7.58**
M-1 vs. average lateral electric field for the strained-Si/SiGe and bulk-Si test structures. After N. S. Waldron et al., *IEEE IEDM Tech. Dig.*, 2003(813-816). With permission.

The higher body current of the strained-Si sample is indicative of higher impact ionization as measured by the impact ionization multiplication coefficient, M-1 = $I_b/I_s$. M-1 is a useful figure of merit to gauge the extra current generated by impact ionization. As seen in Fig. 7.58, for the same average lateral electric field ($E_{avg}$), M-1 in the strained-Si/SiGe devices is close to two orders of magnitude higher than in the bulk-Si sample. Further evidence of enhanced impact ionization in the strained-Si/SiGe samples is the breakdown voltage, which is nearly 4 V lower than in the bulk-Si samples. A positive temperature coefficient for impact ionization has very significant implications in device applications, particularly in RF power applications. A positive temperature coefficient for impact ionization has also been observed in III-V compound semiconductors. In the $In_xGa_{1-x}As$ system, the temperature coefficient of impact ionization changes from being negative for GaAs (x = 0) to becoming positive for InAs (x = 1) at a composition around x = 0.4 [75].

Hot electrons are produced by the high lateral electric field near the drain in short-channel devices. Hot electrons generate electron-hole pairs via impact ionization, which in turn creates more hot carriers. In surface-channel MOSFETs, hot electrons and hot holes are injected into the gate oxide via hot-carrier injection (HCI), leading to the creation of interface states. These interface states cause shifts in the threshold voltage and increased subthreshold slope, which degrades the device performance over time.

While the scaling of power supply voltage has reduced HCI concerns, HCI degradation will still be significant in the deep nanometer regime [76, 77]. When considering high-mobility strained-Si structures, it is clear that the higher mobility should lead to more energetic electrons. Also, the lower bandgap in the strained-Si should make impact ionization easier. These two factors could cause increased HCI. On the other hand, strained-Si should have a higher conduction band offset with $SiO_2$, potentially leading to less HCI.

As experimental evidence suggests that strained-Si/relaxed-$Si_{1-x}Ge_x$ structures have significantly higher levels of impact ionization than unstrained-Si under identical operating conditions [74], it is important to study the hot-electron reliability in MOSFETs in the light of self-heating effect and the high levels of impact ionization. Since a very high longitudinal electric field exists near the drain of short-channel MOSFETs, many electron-hole pairs are created. This provokes a cascading phenomenon in which generated carriers produce even more electron-hole pairs through secondary impact ionization, intensifying the effect. A hot carrier with sufficient energy can produce more charge carriers through impact ionization. For strained-Si, one can expect that the higher carrier saturation velocities that arise from increased channel mobility would enhance the electron-hole generation rate according to Eqn. 7.8. The reduction of the bandgap in strained-Si due to valley splitting would also exacerbate impact ionization effects compared to bulk-Si.

The conventional metric for impact ionization in a MOSFET is the substrate current ($I_{sub}$). For n-MOSFETs, holes generated by impact ionization

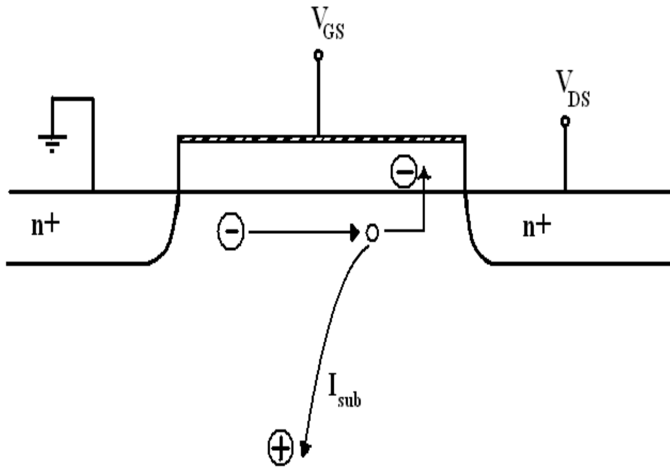



**FIGURE 7.59**
Substrate current as a result of impact ionization near the drain in an n-channel MOSFET.

are collected by the substrate, giving rise to a measurable $I_{sub}$, as depicted in Fig. 7.59. Electrons are either collected by the drain or injected into the gate dielectric. $I_{sub}$ can cause circuit problems, as it leads to noise and CMOS latchup, but it can also be used to indirectly monitor hot-carrier effects. For this reason, the bias conditions that have traditionally been used for performing hot-carrier reliability studies are the biases that correspond to maximum $I_{sub}$. The metric for the hot-carrier degradation rate is $I_{sub}$ normalized by the drain current, or $I_{sub}/I_d$.

### 7.6.1 Self-heating

Self-heating is a pressing issue for both the silicon-on-insulator and strained-Si technologies, where the devices are separated from the silicon substrate by poor thermal conducting layers. Although seemingly counterintuitive, the level of self-heating in a strained-Si transistor could be comparable with that of the SOI device due to the poor thermal conductivity of the thick $Si_{0.8}Ge_{0.2}$ underlayer ($\sim$ 5 W/m-K). The lateral thermal conduction in strained-Si layer of thickness near 10-20 nm would somewhat reduce the maximum temperature rise in the device but is significantly reduced due to the phonon-boundary scattering. In the absence of effective tools for subcontinuum heat transfer modeling, reduced thermal conductivity values for thin silicon and $Si/Si_{0.8}Ge_{0.2}$ are used in a 1D multi-fin model that accounts for the lateral conduction in the channel, source, and drain, as well as heat loss to the $Si_{0.8}Ge_{0.2}$ underlayer. This provides a simple yet effective tool for thermal simulations of the

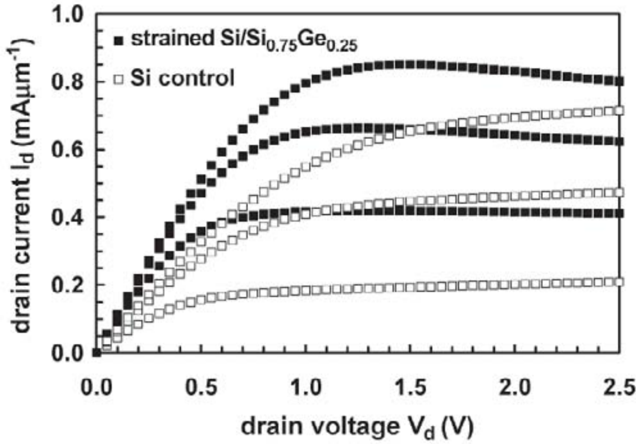



**FIGURE 7.60**
Drain current ($I_d$) vs. drain voltage ($V_d$) characteristics for 0.3 $\mu$m gate length strained-Si/SiGe and Si control devices measured at gate overdrive voltages ($V_g$-$V_t$) of 1, 2, and 3 V showing the self-heating effect. After S. H. Olsen et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(1156-1163). With permission.

strained-Si transistors, which can also be extended to the SiGe-on-insulator (SGOI) technology.

However, the self-heating is a serious concern for both of these devices, which are separated from the silicon substrate by poor thermal conducting layers [78]. The level of self-heating in a strained-Si transistor could he comparable with that of the SOI device due to the poor thermal conductivity ($\sim$ 5 W/m-K) of the thick $Si_{0.8}Ge_{0.2}$ underlayer. The lateral thermal conduction in strained-Si layer of thickness between 30 to 200 Å can somewhat reduce the maximum temperature rise in the device.

In the absence of experimental data and effective tools for subcontinuum heat transfer modeling, reduced thermal conductivity values for thin silicon and $Si/Si_{0.8}Ge_{0.2}$ are used to obtain temperature distribution in a strained-Si transistor. The output characteristics for a short-channel device are shown in Fig. 7.60. The strained-Si device exhibits a strong negative differential output conductance at higher drain voltages. This effect has also been observed in MOSFETs fabricated in silicon-on-isulator materials, having a low thermal conductivity, k $\approx$ 0.014 W/(cm-K). Heat produced in the SOI MOSFETs cannot get dissipated as rapidly as in devices fabricated on bulk-Si. As a result, the device itself heats up above the ambient temperature.

Since mobility is a strong function of temperature, and decreases as the temperature increases, the transconductance of the device also decreases. In

saturation, the drain current is normally constant with decreasing drain bias, or rises slightly due to short-channel effects. However, since the power generated by the transistor is basically $I_d \times V_{ds}$, the result of the mobility degradation caused by self-heating is less drain current as the drain voltage increases. Note that the time associated with self-heating effects in MOSFETs has been found to be on the order of 1 $\mu$s so the integrated measurements performed in msec on the HP4145B parameter analyzer all yield steady state values for the currents.

## 7.7 Industry Example: TSMC

As discussed in chapter 3, strained-Si on relaxed-SiGe substrates or substrate strain technology is a promising candidate for transistor performance enhancement (Fig. 7.61). Strain splits the degeneracy in the conduction and valence bands of Si, enhances the transport properties of electrons and holes, and provides transistor speed enhancement and, hence, power reduction. However, current substrate strained-Si technology faces several process integration challenges that limit its manufacturability. For example, all substrate strained-Si transistors reported to date have been mostly laboratory demonstration having narrow-width (W) transistors due to the low yield of wide-channel transistors. Substrate strained-Si process needs to be optimized to enhance manufacturability and circumvent difficulties associated with the integration of the strained-Si/SiGe heterostructure.

In the following, we have chosen a manufacturable substrate strained-Si technology reported by TSMC [26]. A ring oscillator speed of 6.5 ps at 1.2 V operation for sub-65 nm substrate strained-Si transistors, and process window optimization for the integration of a Si/SiGe have been reported. An enhancement (15%) in $I_{on}$-$I_{off}$ characteristics without correction for self-heating effect was also reported. The transistors featuring a 1.7-nm gate oxide were fabricated on Si and Si/SiGe substrates.

In long-channel devices, electron mobility enhancement of up to 80% was obtained with 20% Ge (Fig. 7.62). To fairly evaluate the benefit of strained-Si on short-channel transistors, the $V_t$ roll-off characteristics of strained-Si and control devices were exactly matched (Fig. 7.63). At a fixed off-state leakage, the nMOS drive current $I_{on}$ is enhanced by 15% (Fig. 7.64) without correcting for self-heating. The I-V characteristics of 80-nm $L_g$ strained-Si and control transistors are presented in Figs. 7.65 and 7.66. For comparable subthreshold slopes, DIBL, and $I_{off}$ (Fig. 7.65), 18% current drive enhancement at $V_{dd}$ = 1.0 V is clearly seen in the $I_{ds}$-$V_{ds}$ curves (Fig. 7.66) despite the lower output conductance of strained-Si transistors due to self-heating.

Fig. 7.67 shows that short-channel effects of strained-Si transistors can be

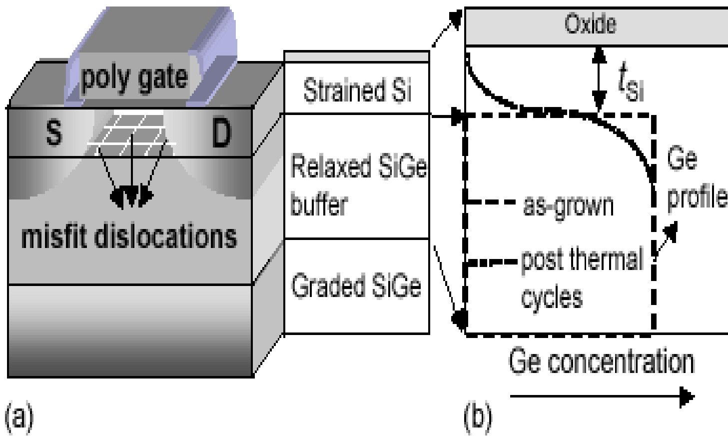



**FIGURE 7.61**
Schematic illustration of a strained-Si transistor, (a) misfit dislocations at the interface between strained-Si layer and relaxed-SiGe layer result from partial relaxation. (b) Ge atoms diffuse from relaxed-SiGe layer to the surface. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

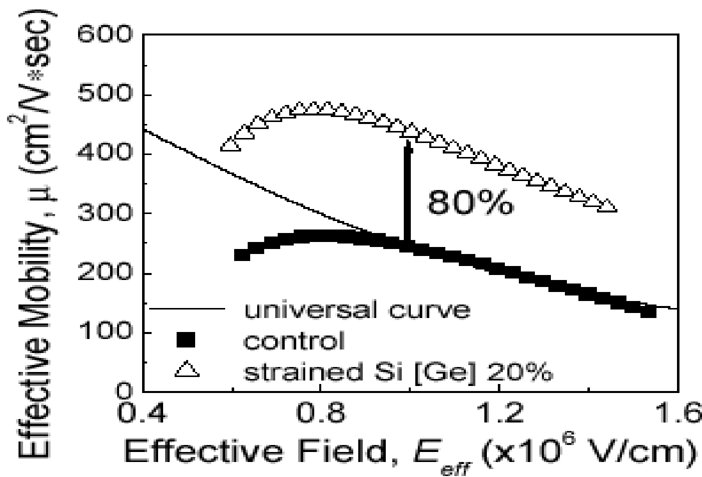



**FIGURE 7.62**
Strained-Si transistor exhibits 80% electron mobility enhancement at $E_{eff}$ = 1 MV/cm. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

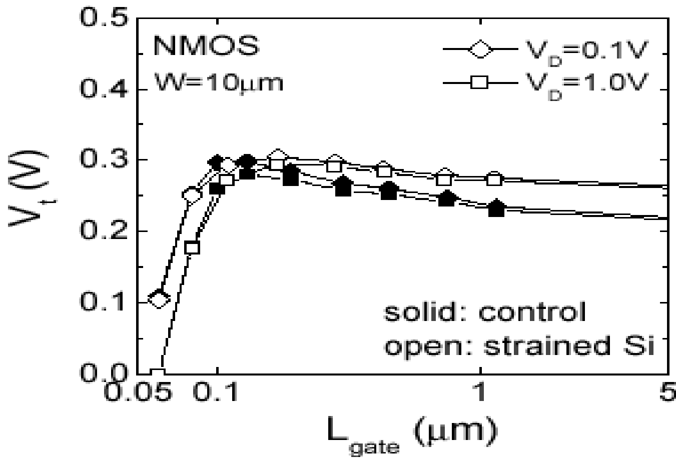



**FIGURE 7.63**
$V_t$ roll-off characteristics of strained Si and control transistors were matched exactly. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

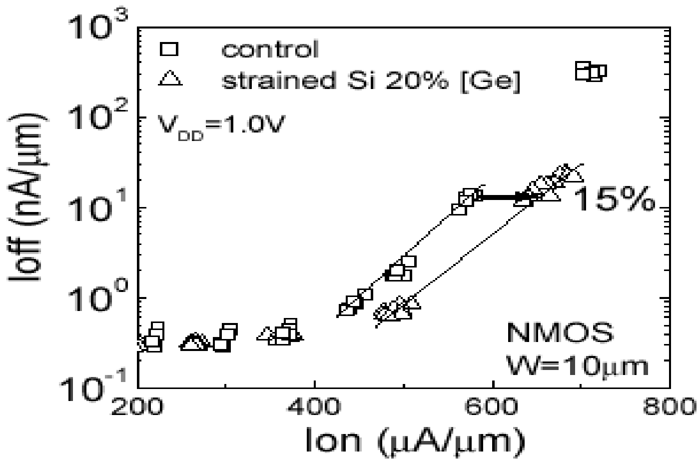



**FIGURE 7.64**
15% enhancement is the largest ever reported for strained-Si transistors. Correction of self-heating effect adds 10% more enhancement. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

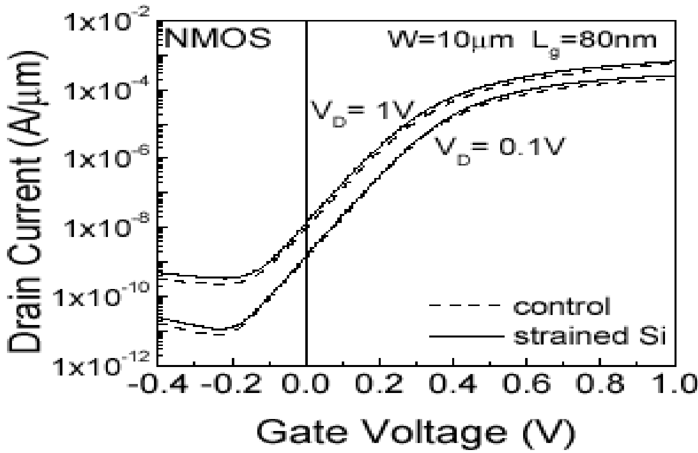



**FIGURE 7.65**
Comparable $I_{ds}$-$V_{gs}$ characteristics of 80-nm $L_g$ strained-Si and control transistors. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

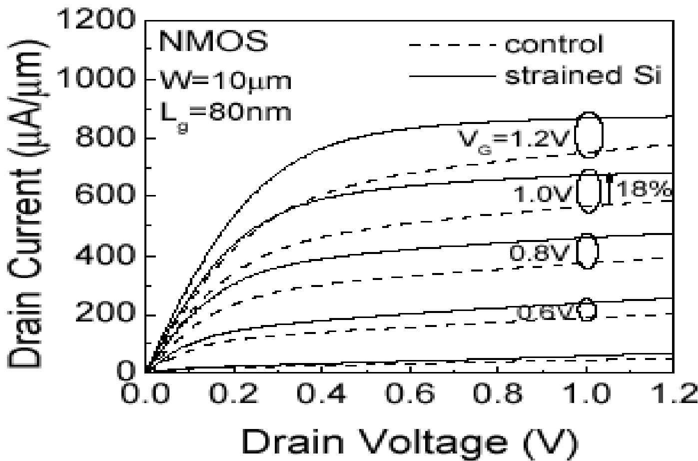



**FIGURE 7.66**
$I_{ds}$-$V_{ds}$ characteristics of 80-nm $L_g$ strained-Si and control transistors. 18% $I_{on}$ improvement is achieved at $V_{dd} = 1.0$ V. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

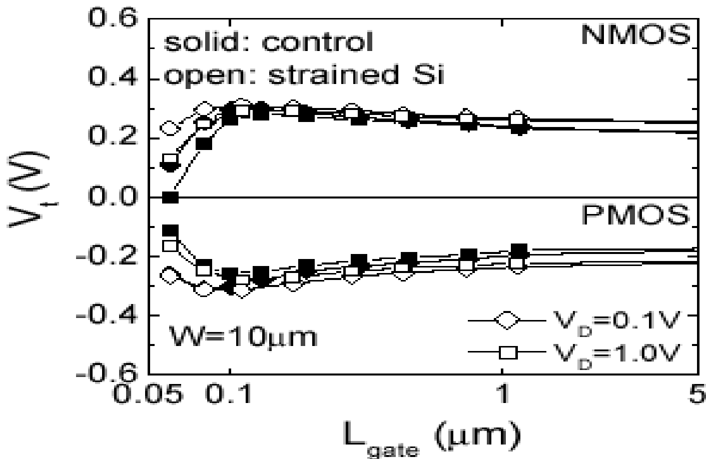



**FIGURE 7.67**
$V_t$ roll-off characteristics of strained-Si and control transistors. Good SCE control is observed down to 60 nm $L_g$ for strained-Si transistors. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

well controlled down to $L_g$ = 60 nm despite the enhanced diffusion of As and P in SiGe. Figs. 7.68 and 7.69 show the I-V characteristics of 60 nm $L_g$ strained-Si transistors. It should be pointed out that optimizing halo and source/drain implants to minimize the junction capacitance $C_J$ is crucial for realizing the advantages of strained-Si. $C_J$ can be deteriorated severely for nonoptimized strained-Si transistors, negating the expected speed improvement resulting from current drive enhancement. The same or slightly lower $C_J$ can be achieved for strained-Si transistors by optimizing halo and source/drain implants. A speed of 6.5 ps for the unloaded ring oscillator is obtained at $V_{dd}$ = 1.2 V, the smallest value reported for strained-Si transistors (Fig. 7.70).

A key integration issue is defect reduction. Width dependence of short-channel effect in strained-Si transistors has been studied. A comparison of the $V_t$ roll-off characteristics for wide- and narrow-channel transistors demonstrates that wide-channel strained-Si transistors are more susceptible to $V_t$ roll-off (Fig. 7.71). Wide- and long-channel (W/L = 10/10 $\mu$m) strained-Si transistors reveal the characteristic orthogonal pattern of misfit dislocations due to relaxation of the Si/SiGe interface. For a given degree of strain relaxation and interfacial dislocation density, narrow-width transistors are less likely to intersect a dislocation and suffer from increased off-state leakage. This is the reason for the low yield of wide-channel strained-Si transistors and the SCE width dependence as shown in Fig. 7.71.

It has been shown that by reducing the thickness of the strained-Si layer $t_{Si}$ below a value $t_{Si,max}$, width-dependent SCE can be effectively suppressed.

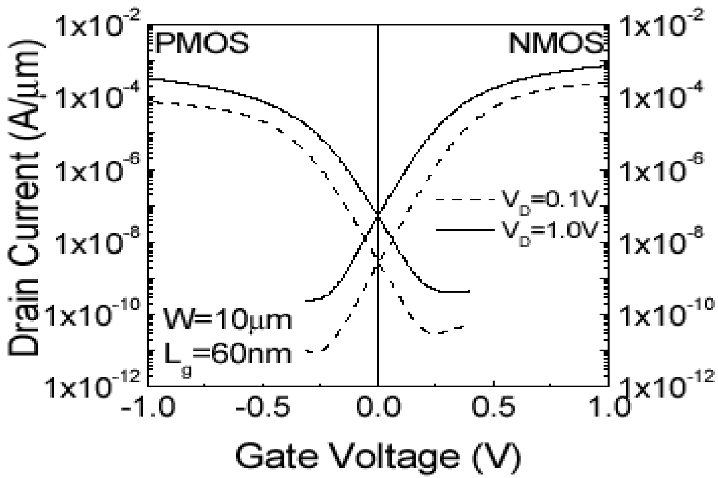



**FIGURE 7.68**
$I_{ds}$-$V_{gs}$ characteristics of 60-nm $L_g$ strained-Si transistors. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

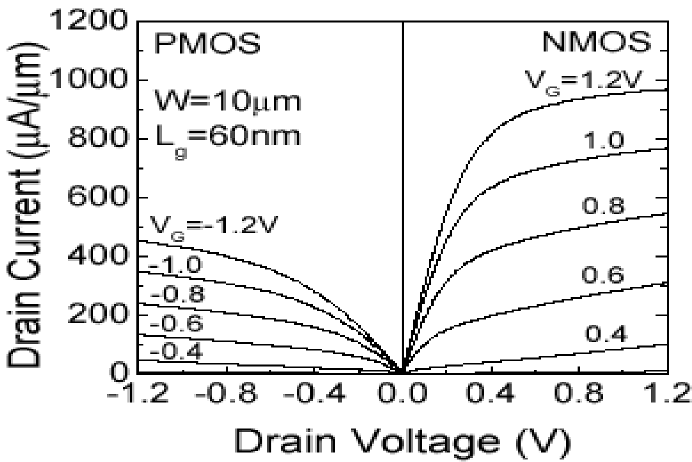



**FIGURE 7.69**
$I_{ds}$-$V_{ds}$ characteristics of 60-nm $L_g$ strained-Si transistors. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

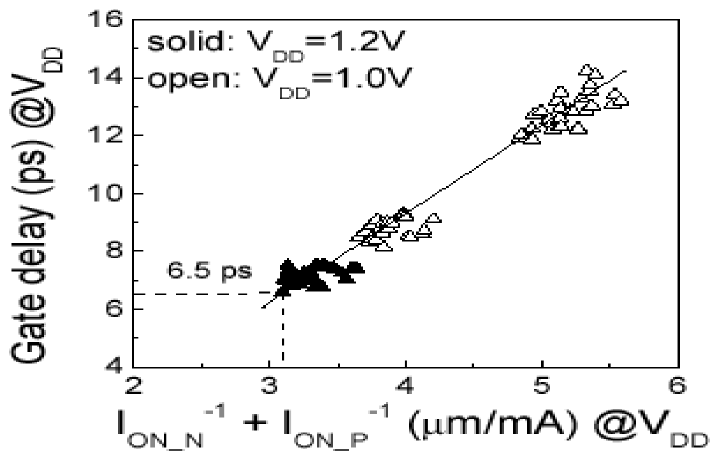



**FIGURE 7.70**
Unloaded strained-Si ring oscillator gate delay. A speed of 6.5 ps is obtained at $V_{dd} = 1.2$ V. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

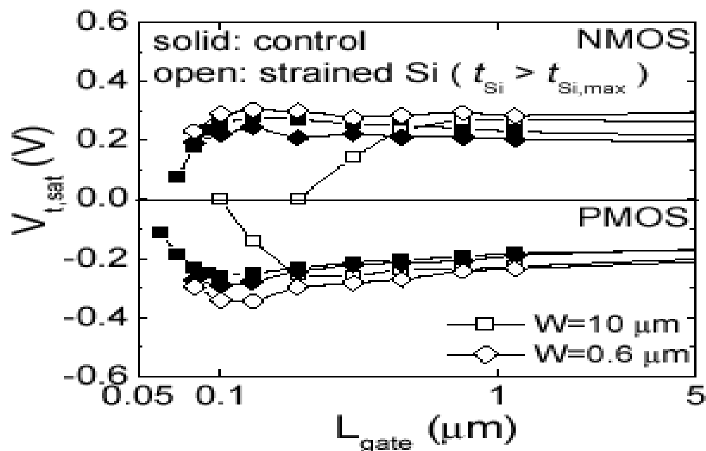



**FIGURE 7.71**
SCE is significantly degraded in wide-channel strained-Si transistors. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

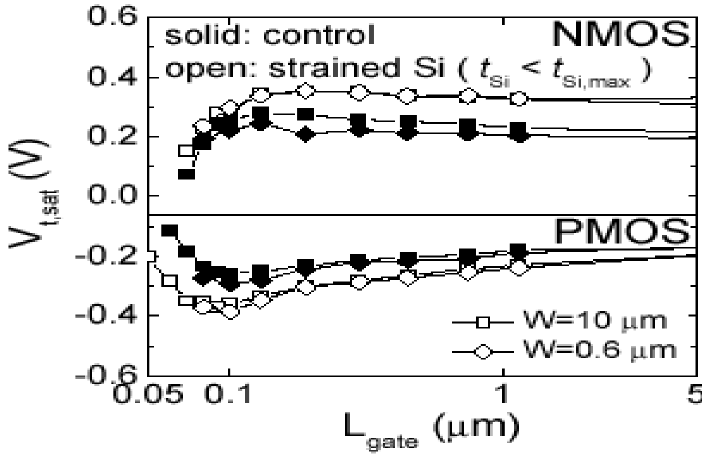



**FIGURE 7.72**
Short-channel effect: width dependence in strained-Si transistors is suppressed by $t_{Si} < t_{Si,max}$. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

Strain relaxation occurs for $t_{Si}$ greater than $t_{Si,max}$, so that thinner $t_{Si}$ can eliminate misfit dislocations altogether. The $V_t$ roll-off characteristics for wide-channel strained-Si transistors are thus improved (Fig. 7.72). This enables the fabrication of wide-channel strained-Si transistors.

Due to critical layer thickness limitation, it is known that a higher Ge mole fraction in the relaxed-SiGe layer requires thinner $t_{Si,max}$, consistent with the model of defect introduction via partial relaxation. However, a reduction of $t_{Si}$ and/or increase in the Ge concentration in the substrate increases the likelihood of Ge atoms diffusing to the interface between the gate dielectric and the strained-Si channel [16]. This introduces interface states and oxide charge as revealed in the C-V characteristics (Fig. 7.73) and degrades the carrier mobility (Fig. 7.74). Mobility degradation concerns require that $t_{Si}$ be kept above a minimum value $t_{Si,min}$. Evaluating the effect of Ge concentration on $t_{Si,min}$, we note that $t_{Si,min}$ increases with increasing Ge concentration (Fig. 7.75).

To prevent the partial relaxation-induced misfit dislocations, $t_{Si}$ should be kept below $t_{Si,max}$. On the other hand, $t_{Si}$ must be greater than $t_{Si,min}$ to avoid mobility degradation by interface states and oxide charge arising from Ge diffusion to the surface. Therefore, a definition for the process window for $t_{Si}$ is shown in Fig. 7.76. It is worth noting that the process window narrows with increasing thermal budget since both strain relaxation and Ge diffusion are thermally activated. Choosing an optimal process window is important for obtaining a robust and manufacturable substrate strained-Si technology.

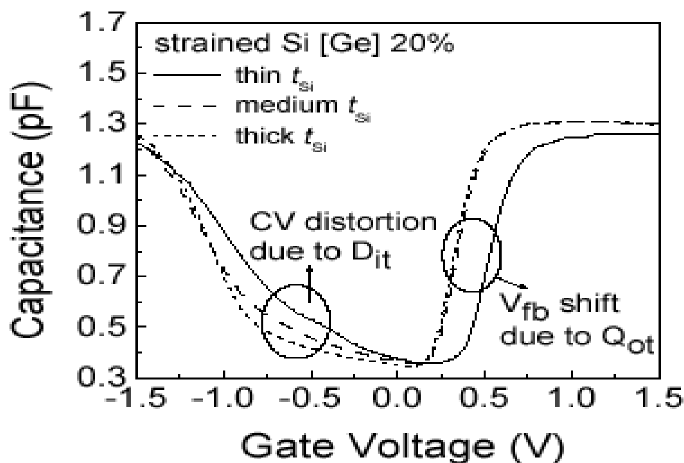



**FIGURE 7.73**
C-V characteristics of strained-Si n-MOSFETs. Interface states and oxide charge are found for strained-Si transistors with thin $t_{Si}$. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

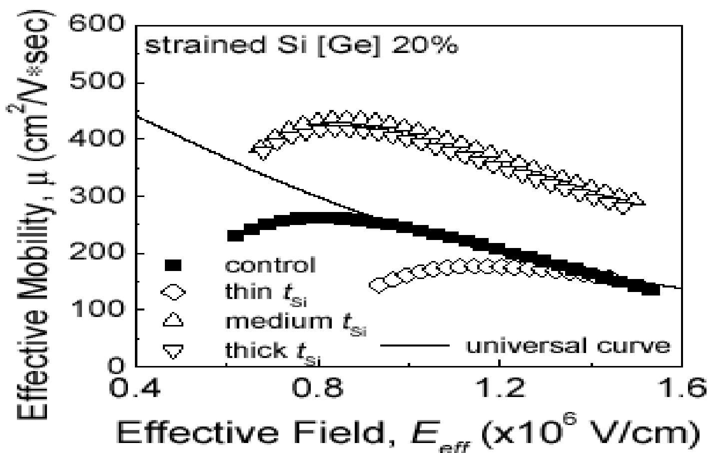



**FIGURE 7.74**
Effect of $t_{Si}$ on electron mobility, showing degradation for thin $t_{Si}$. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

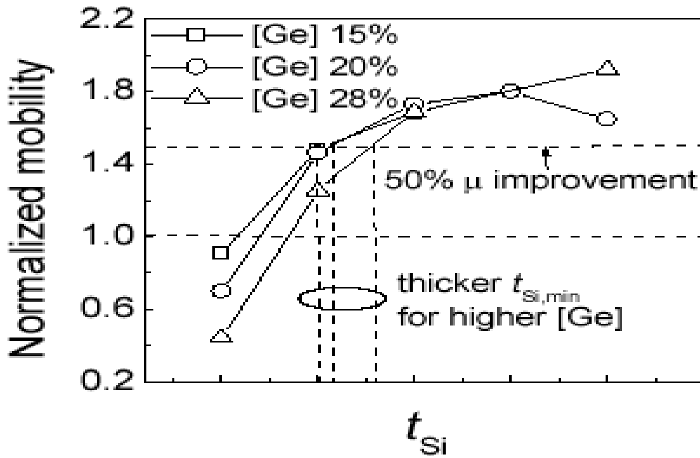



**FIGURE 7.75**
Effect of Ge concentration on $t_{Si,min}$ for strained-Si transistors. Thicker $t_{Si,min}$ is required for higher Ge concentration. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

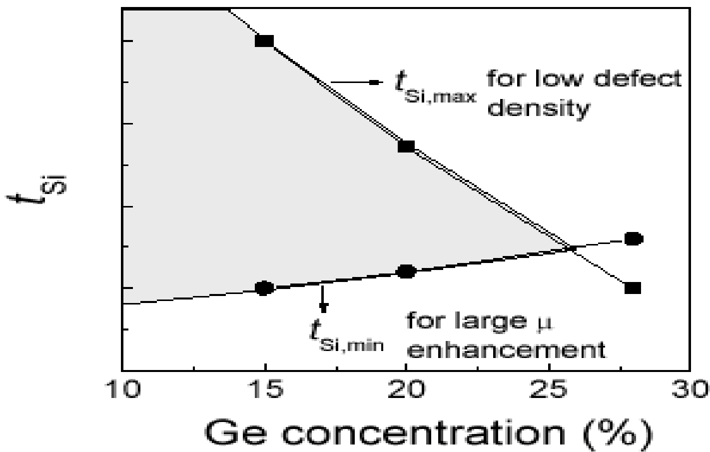



**FIGURE 7.76**
Process window for strained-Si transistors. Selection of $t_{Si}$ for a given thermal budget is limited by the constraints related to misfit defect control and mobility enhancement. After H. C.-H. Wang et al., *IEEE IEDM Tech. Dig.*, 2003(61-64). With permission.

## 7.8 Industry Example: AMD

Transistor scaling has been the primary factor for driving CMOS performance improvement for more than last 40 years. Toward further scaling, strained-Si is being considered to enhance FET performance. In addition to conventional scaling challenges, key issues for strained-Si devices include Si/SiGe band-offset induced $V_t$ shift, differential diffusion of dopants in SiGe, limited strained-Si thickness, strain relaxation, and Ge diffusion at high temperatures. Xiang et al. [14] have reported on the performance enhancement and scalability for strained-Si CMOS devices with $L_g$ down to 25 nm, 1.2 nm nitrided oxide, and NiSi. Good control of short-channel effects was achieved. NiSi was found to be superior to $CoSi_2$ for strained-Si devices. Electron mobility enhancement resulted in significant $I_{dsat}$ improvement for long channel nMOS. Due to high doping and significant DC self-heating, the $I_{dsat}$ improvement diminishes with gate length scaling. Fig. 7.77 shows a strained-Si MOSFET with a 25-nm $L_g$ patterned by 193-nm lithography. Figs. 7.78 and 7.79 show the transistor's I-V characteristics. In experiment, the drive currents of $\sim 615$ $\mu$A/$\mu$m (nMOS) $\sim 317$ $\mu$A/$\mu$m (pMOS) are measured at $V_{dd}$ of 0.9 V, with $I_{off} < 100$ nA/$\mu$m for both devices.

With thin gate oxide and optimized $V_t$/halo implants, good control of short-channel effects was achieved for both p- and nMOS, despite enhanced As diffusion in nMOS. Fig. 7.80 shows good threshold voltage roll-off characteristics down to 25-nm gate length. Silicide was found to be a key module for realization of the strained-Si technology. SiGe in strained-Si substrate inhibits Co silicide transition to the low resistivity phase, resulting in a high silicide sheet resistance with a large variation. Ni reacts with both strained-Si and SiGe and forms low sheet resistance silicide and germanosilicide [17]. Fig. 7.81 shows the sheet resistance and standard deviation for Ni and Co silicides in strained-Si devices with typical silicidation processes.

Measured electron and hole mobility characteristics are shown in Fig. 7.82. Compared to the unstrained device, strained-Si nMOS displays significant mobility enhancements over a wide $E_{eff}$ range. No significant hole mobility improvement was observed for the strained-Si structure studied. The electron mobility enhancement results in significant $I_{dsat}$ improvement ($> 50\%$) for long-channel nMOS (Fig. 7.83). However, the $I_{dsat}$ improvement diminishes as the gate length is scaled down. There is almost no $I_{dsat}$ improvement for 25-nm devices at room temperature. The reduction in $I_{dsat}$ enhancement is believed to be due to high channel doping and severe self-heating in short-channel devices. The high channel doping in short-channel devices caused by halo implants degrades mobility. Low thermal conductivity of SiGe and high drive current of short-channel nMOS cause severe self-heating during the DC measurements [2]. However, this is not expected to have a serious impact on transistor performance for digital applications, since the thermal time con-

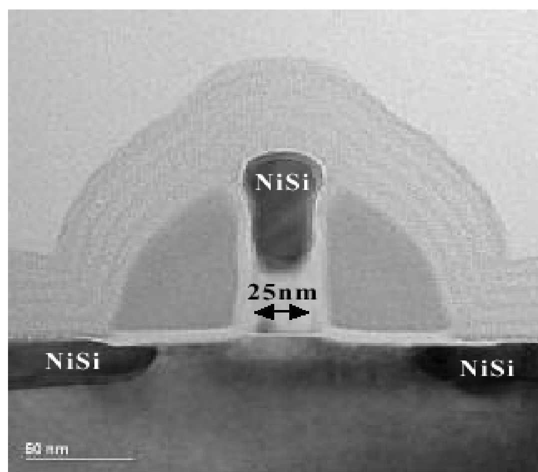



**FIGURE 7.77**
TEM cross-section of a 25-nm physical gate length strained-Si MOS transistor. After Q. Xiang et al., *Proc. ISTDM*, 2003.

stant for degradation (greater than 100 ns) is much larger than the typical transistor on time [79]. At a high temperature, influences of self-heating and Coulomb scattering decreased and greater improvement was achieved. Despite similar $L_{gate}$ and temperature dependence, more $I_{dlin}$ improvement was observed due to an increasing influence of mobility and decreasing self-heating effects (Fig. 7.84). Consistently, greater improvement in ring oscillator speed was observed for a low $V_{dd}$ at a high temperature (see Fig. 7.85).

## 7.9 Summary

"First generation" global and local stress techniques are progressing rapidly. In this chapter, strained-Si CMOS technology and MOSFET performance enhancement by using strained-Si have been discussed. In this chapter, we have examined various MOSFET channel designs in the SiGe/Relaxed-SiGe/Si system to explore the limits of electron and hole mobility in inversion layers. It is observed that the main challenge in enhancing MOSFET performance beyond conventional strained-Si lies in understanding structures that enhance hole mobility at high inversion charge densities. Various stress used in leading-edge advanced CMOS has been discussed in depth, along with simulation examples. The emphasis for this work was to highlight the issues which the industry needs to account for in developing the process and devices.

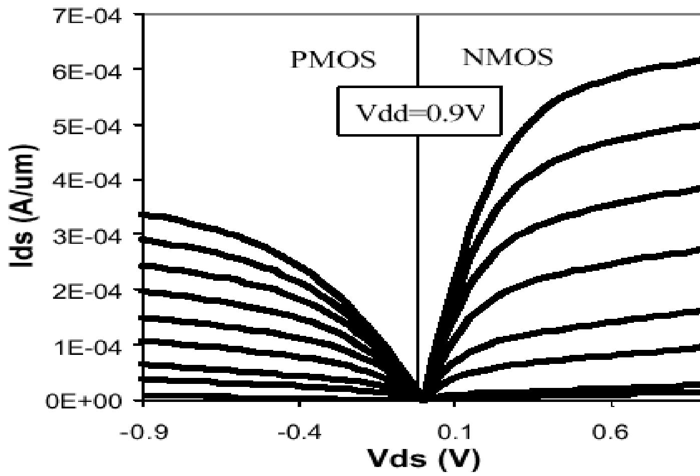



**FIGURE 7.78**
25-nm strained-Si CMOS $I_{ds}$-$V_{ds}$ characteristics. After Q. Xiang et al., *Proc. ISTDM*, 2003.

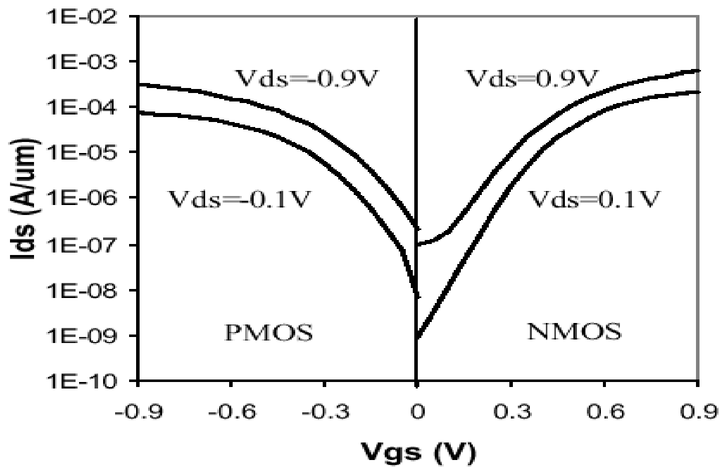



**FIGURE 7.79**
25-nm strained-Si CMOS $I_{ds}$-$V_{gs}$ characteristics. After Q. Xiang et al., *Proc. ISTDM*, 2003.

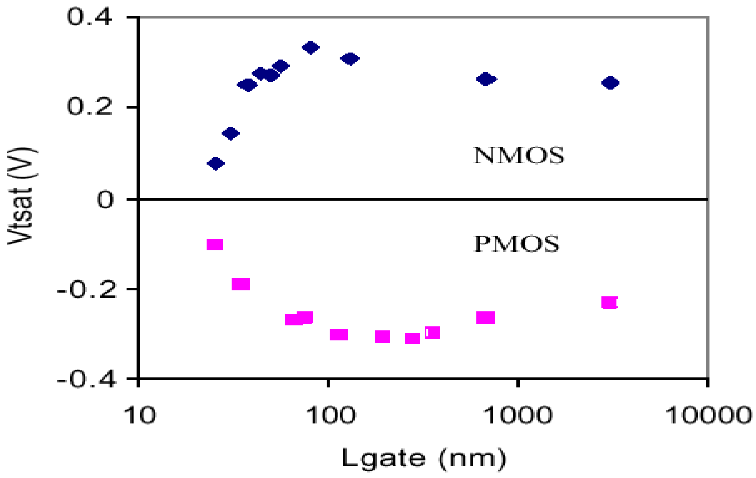



**FIGURE 7.80**
Threshold voltage roll-off of strained-Si CMOS. After Q. Xiang et al., *Proc. ISTDM*, 2003.

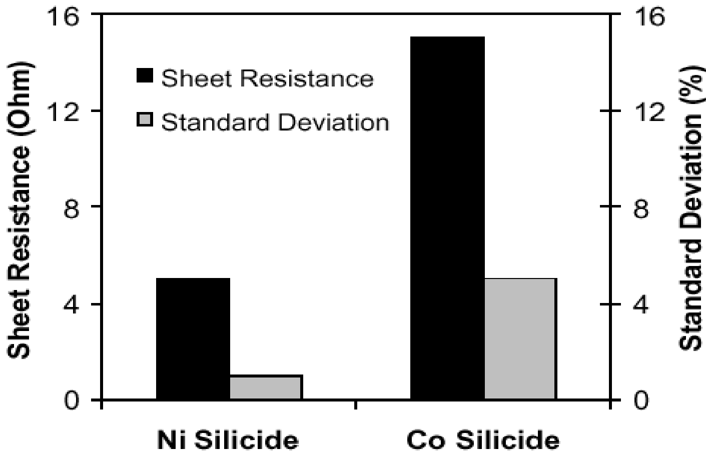



**FIGURE 7.81**
Sheet resistance and standard deviation of Ni and Co silicide for strained-Si devices. After Q. Xiang et al., *Proc. ISTDM*, 2003.

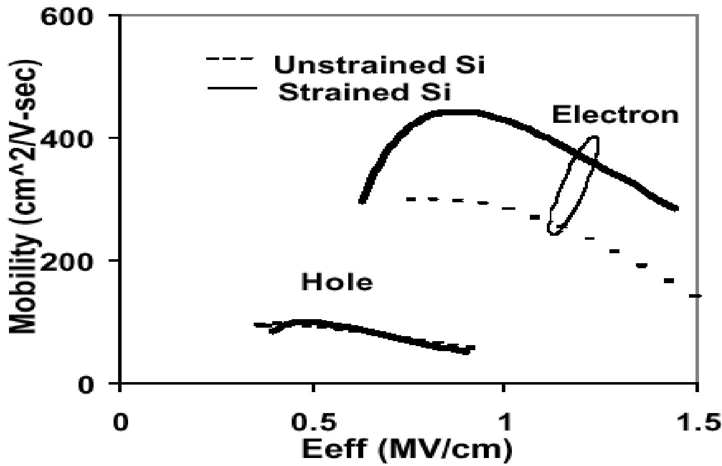



**FIGURE 7.82**
Effective mobilities of electrons (nMOS) and holes (pMOS) in strained-Si. After Q. Xiang et al., *Proc. ISTDM*, 2003.

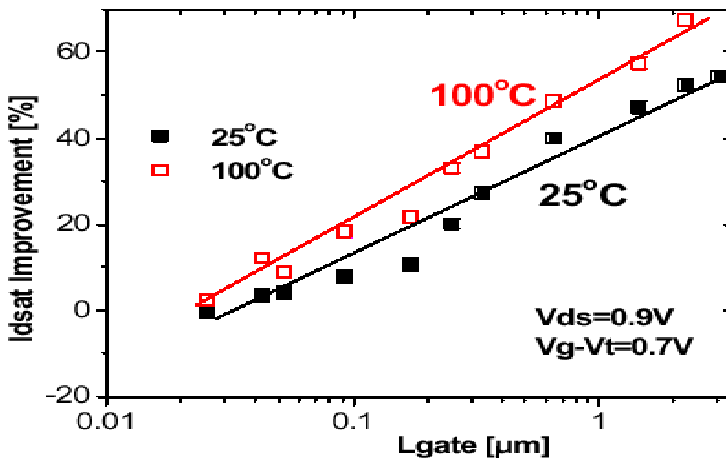



**FIGURE 7.83**
Percentage saturation drive current enhancement for strained-Si nMOS as compared to unstrained-Si nMOS control. Data from K. Rim et al., *VLSI Tech. Dig.*, p. 98, 2002 is pulsed measurement to eliminate self-heating. After Q. Xiang et al., *Proc. ISTDM*, 2003.

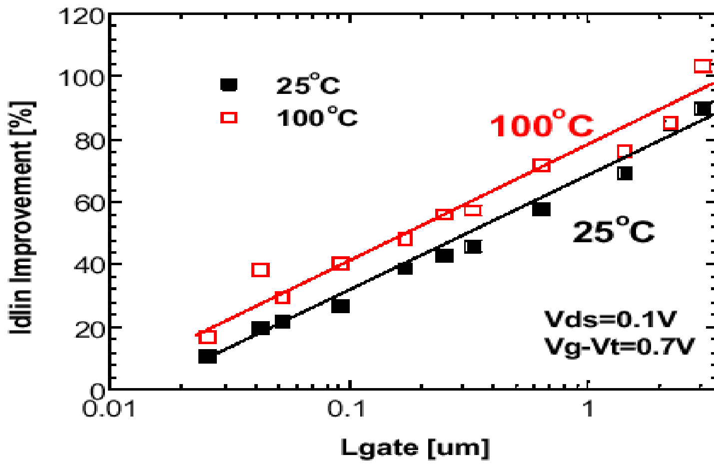



**FIGURE 7.84**
Percentage linear drive current enhancement for strained-Si nMOS as compared to unstrained-Si nMOS control. After Q. Xiang et al., *Proc. ISTDM*, 2003.

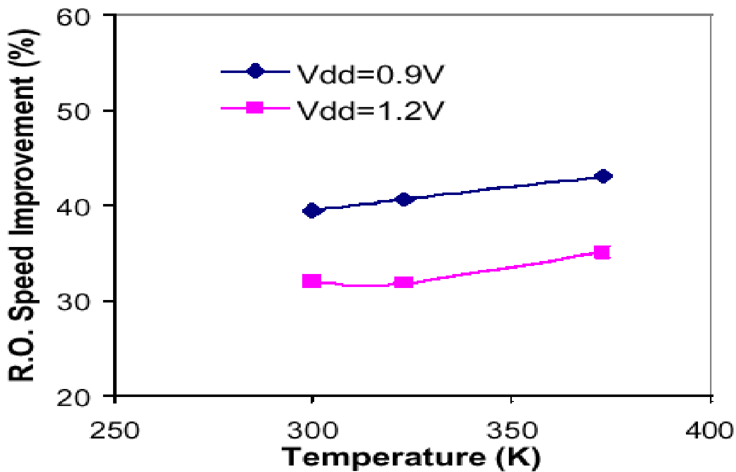



**FIGURE 7.85**
Improvement of ring oscillator of 35-nm strained-Si CMOS as compared to unstrained-Si CMOS control. After Q. Xiang et al., *Proc. ISTDM*, 2003.

Understanding the magnitude of the threshold voltage shift is important when determining the performance gain of strained-Si. The experimental data suggest large threshold voltage shifts occur for biaxial stress, and much smaller shifts under uniaxial stress. Based on metrics for cost, process complexity, magnitude of the performance gain, and scalability, strained-Si is the clear choice for extending CMOS when simple scaling ended at the 90-nm technology node. Probably, with limited redesign to take full advantage of the transistor enhancements, product roadmaps can be extended even further with strained-Si.

## References

[1] N. D. Jankovic and A. O'Neill, "Enhanced performance virtual substrate heterojunction bipolar transistor using strained-Si/SiGe emitter," *Semicond. Sci. Technol.*, vol. 18, pp. 901–906, 2003.

[2] K. Rim, J. Chu, H. Chen, K. A. Jenkins, T. Kanarsky, K. Lee, A. Mocuta, H. Zhu, R. Roy, J. Newbury, J. Ott, K. Petrarca, P. Mooney, D. Lacey, S. Koester, K. Chan, D. Boyd, M. Leong, and H.-S. Wong, "Characteristics and device design of sub-100 nm strained-Si n- and p-MOSFETs," *IEEE VLSI Tech. Symp. Dig.*, pp. 98–99, 2002.

[3] D. K. Nayak, J. C. S. Woo, J. S. Park, K. L. Wang, and K. P. MacWilliams, "High-mobility p-channel metal-oxide-semiconductor field-effect transistor on strained-Si," *Appl. Phys. Lett.*, vol. 62, pp. 2853–2855, 1993.

[4] R. W. Keyes, "High-mobility FET in strained silicon," *IEEE Trans. Electron Dev.*, vol. ED-33, p. 863, 1986.

[5] J. Welser, J. L. Hoyt, and J. F. Gibbons, "NMOS and PMOS transistor fabricated in strained silicon/relaxed silicon-germanium structures," *IEEE IEDM Tech. Dig.*, pp. 1000–1003, 1992.

[6] N. Mohta and S. E. Thompson, "Mobility enhancement - the next vector to extend Moore's law," *IEEE Circuits and Dev. Mag.*, vol. September/October, pp. 18–23, 2005.

[7] W. Zhang and J. G. Fossum, "On the threshold voltage of strained-Si-$Si_{1-x}Ge_x$ MOSFETs," *IEEE Trans. Electron Dev.*, vol. 52, pp. 263–268, 2005.

[8] J.-S. Lim, S. E. Thompson, and J. G. Fossum, "Comparison of threshold-voltage shifts for uniaxial and biaxial tensile-stressed n-MOSFETs," *IEEE Electron Dev. Lett.*, vol. 25, pp. 731–733, 2004.

[9] S. E. Thompson, M. Armstrong, C. Auth, S. Cea, R. Chau, G. Glass, T. Hoffman, J. Klaus, Z. Ma, B. Mcintyre, A. Murthy, B. Obradovic, L. Shifren, S. Sivakumar, S. Tyagi, T. Ghani, K. Mistry, M. Bohr, and Y. El-Mansy, "A logic nanotechnology featuring strained-silicon," *IEEE Electron Dev. Lett.*, vol. 25, pp. 191–193, 2004.

[10] S. E. Thompson, G. Sun, K. Wu, J. Kim, and T. Nishida, "Key differences for process-induced uniaxial vs. substrate-induced biaxial stressed Si and Ge channel MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 221–224, 2004.

[11] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, "Hole mobility enhancements and alloy scattering-limited mobility in tensile strained-Si/SiGe surface channel metal-oxide-semiconductor field-effect transistors," *J. Appl. Phys.*, vol. 92, pp. 3745–3751, 2002.

[12] M. V. Fischetti, Z. Ren, P. M. Solomon, M. Yang, and K. Rim, "Six-band k·p calculation of the hole mobility in silicon inversion layers: dependence on surface orientation, strain, and silicon thickness," *J. Appl. Phys.*, vol. 94, pp. 1079–1095, 2003.

[13] J.-S. Goo, Q. Xiang, Y. Takamura, F. Arasnia, E. N. Paton, P. Besser, J. Pan, and M.-R. Lin, "Band offset induced threshold variation in strained-Si nMOSFETs," *IEEE Electron Dev. Lett.*, vol. 24, pp. 568–570, 2003.

[14] Q. Xiang, J.-S. Goo, J. Pan, B. Yu, S. Ahmed, J. Zhang, and M.-R. Lin, "Strained silicon NMOS with nickel-silicide metal gate," *IEEE VLSI Tech. Symp. Dig.*, pp. 101–102, 2003.

[15] N. Sugii, D. Hisamoto, K. Washio, N. Yokoyama, and S. Takagi, "Performance enhancement of strained-Si MOSFETs fabricated on a chemical-mechanical-polished SiGe substrate," *IEEE Trans. Electron Dev.*, vol. 49, pp. 2237–2243, 2002.

[16] M. T. Currie, C. W. Leitz, T. A. Langdo, G. Taraschi, E. A. Fitzgerald, and D. A. Antoniadis, "Carrier mobilities and process stability of strained-Si n- and p-MOSFETs on SiGe virtual substrates," *J. Vac. Sci. Technol. B*, vol. 19, pp. 2268–2279, 2001.

[17] H. B. Zhao, K. L. Pey, W. K. Choi, S. Chattopadhyay, E. A. Fitzgerald, D. A. Antoniadis, and P. S. Lee, "Interfacial reactions of Ni on $Si_{1-x}Ge_x$ (x = 0.2, 0.3) at low temperature by rapid thermal annealing," *J. Appl. Phys.*, vol. 92, pp. 214–217, 2002.

[18] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K.

Zawadzki, S. Thompson, and M. Bohr, "A 90 nm high volume manufacturing logic technology featuring novel 45 nm gate length strained silicon CMOS," *IEEE IEDM Tech. Dig.*, pp. 978–980, 2003.

[19] K. W. Ang, K. J. Chui, V. Bliznetsov, A. Du, N. Balasubramanian, M. F. Li, G. Samudra, and Y.-C. Yeo, "Enhanced performance in 50 nm N-MOSFETs with silicon-carbon source/drain regions," *IEEE IEDM Tech. Dig.*, pp. 1069–1071, 2004.

[20] K.-W. Ang, K.-J. Chui V. Bliznetsov, C.-H. Tung, A. Du, N. Balasubramanian, G. Samudra, M. F. Li, and Y.-C. Yeo, "Lattice strain analysis of transistor structures with silicon-germanium and silicon-carbon source/drain stressors," *Appl. Phys. Lett.*, vol. 86, pp. 093102–1–093102–3, 2005.

[21] M. E. Law, "FLOOPS 2002 Release Manual," 2002.

[22] DESSIS-ISE, Release 8.5, "Integrated Systems Engineering AG, Zurich, Switzerland," 2003.

[23] J. L. Egley and D. Chidambarrao, "Strain effects on device characteristics: implementation in drift-diffusion simulators," *Solid-State Electron.*, vol. 36, pp. 1653–1664, 1993.

[24] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Transconductance enhancement in deep submicron strained-Si n-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 707–710, 1998.

[25] K. Rim, S. Koester, M. Hargrove, J. Chu, P. M. Mooney, J. Ott, T. Kanarsky, P. Ronsheim, M. Ieong, A. Grill, and H.-S. P. Wong, "Strained-Si NMOSFETs for high performance CMOS technology," *VLSI Technology*, pp. 59–60, 2001.

[26] H. C.-H. Wang, Y.-P. Wang, S.-J. Chen, C.-H. Ge, S. M. Ting, J.-Y. Kung, R.-L. Hwang, H.-K. Chiu, L. C. Sheu, P.-Y. Tsai, L.-G. Yao, S.-C. Chen, H.-J. Tao, Y.-C. Yeo, W.-C. Lee, and C. Hu, "Substrate-strained silicon technology: process integration," *IEEE IEDM Tech. Dig.*, pp. 61–64, 2003.

[27] T. Sanuki, A. Oishi, Y. Morimasa, S. Aota, T. Kinoshita, R. Hasumi, Y. Takegawa, K. Isobe, H. Yoshimura, M. Iwai, K. Sunouchi, and T. Noguchi, "Scalability of strained silicon CMOSFET and high drive current enhancement in the 40 nm gate length technology," *IEEE IEDM Tech. Dig.*, pp. 65–68, 2003.

[28] J. W. Matthews and A. E. Blakeslee, "Defects in epitaxial multilayers - I. misfit dislocations," *J. Crystal Growth*, vol. 27, pp. 118–125, 1974.

[29] C. Detavernier, A. Ozcan, C. Lavoie, J.-J. Sweet, and J. M. E. Harper, "Kinetics of agglomeration of NiSi and $NiSi_2$ phase formation," *Proc. Mat. Res. Soc.*, vol. 745, p. N4.12, 2003.

[30] E. Augendre, R. Rooyackers, M. Caymax, E. P. Vandamme, A. De Keersgieter, C. Perello, M. van Dievel, S. Pochet, and G. Badenes, "Elevated source/drain by sacrificial selective epitaxy for high performance deep submicron CMOS: process window versus complexity," *IEEE Trans. Electron Dev.*, vol. 47, pp. 1484–1491, 2000.

[31] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Fabrication and analysis of deep submicron strained-Si N-MOSFETs," *IEEE Trans. Electron Dev.*, vol. 47, pp. 1406–1415, 2000.

[32] T. Morimoto, T. Ohguro, H. S. Momose, T. Iinuma, I. Kunishima, K. Suguro, I. Katakabe, H. Nakajima, M. Tsuchiaki, M. Ono, Y. Katsumata, and H. Iwai, "Self-aligned nickel-mono-silicide technology for high-speed deep submicrometer logic CMOS ULSI," *IEEE Trans. Electron Dev.*, vol. 42, pp. 915–922, 1995.

[33] K. L. Pey, W. K. Choi, S. Chattopadhyay, H. B. Zhao, E. A. Fitzgerald, D. A. Antoniadis, and P. S. Lee, "Thermal reaction of nickel and $Si_{0.75}Ge_{0.25}$ alloy," *J. Vac. Sci. Technol. A*, vol. 20, pp. 1903–1910, 2002.

[34] T. Jarmar, J. Seger, F. Ericson, D. Mangelinck, U. Smith, and S. L. Zhang, "Morphological and phase stability of nickel-germanosilicide on $Si_{1-x}Ge_x$ under thermal stress," *J. Appl. Phys.*, vol. 92, pp. 7193–7199, 2002.

[35] T. Mizuno, N. Sugiyama, T. Tezuka, T. Numata, and S. Takagi, "Thin-film strained-SOI CMOS devices - physical mechanisms for reduction of carrier mobility," *IEEE Trans. Electron Dev.*, vol. 51, pp. 1114–1121, 2004.

[36] T. Numata, K. Uchida, J. Koga, and S. Takagi, "Device design for subthreshold slope and threshold voltage control in sub-100 nm fully-depleted SOI MOSFETs," *IEEE SOI Conf.*, pp. 179–180, 2002.

[37] T. Mizuno, N. Sugiyama, T. Tezuka, T. Numata, and S. Takagi, "High performance CMOS operation of strained-SOI MOSFETs using thin film SiGe-on-insulator," *IEEE VLSI Tech. Symp. Dig.*, pp. 106–109, 2002.

[38] C.-H. Choi, J.-H. Chun, and R. W. Dutton, "Electrothermal characteristics of strained-Si MOSFETs in high-current operation," *IEEE Trans. Electron Dev.*, vol. 51, pp. 1928–1931, 2004.

[39] G. Scott, J. Lutze, M. Rubin, F. Nouri, and M. Manley, "NMOS drive current reduction caused by transistor layout and trench isolation induced stress," *IEEE IEDM Tech. Dig.*, pp. 827–830, 1999.

[40] A. Steegen, M. Stucchi, A. Lauwers, and K. Maex, "Silicide induced pattern density and orientation dependent transconductance in MOS transistor," *IEEE IEDM Tech. Dig.*, pp. 497–500, 1999.

[41] F. Ootsuka, S. Wakahara, K. Ichinose, A. Honzawa, S. Wada, H. Sato, T. Ando, H. Ohta, K. Watanabe, and T. Onai, "A highly dense, high-performance 130 nm node CMOS technology for large scale system-on-a-chip applications," *IEEE IEDM Tech. Dig.*, pp. 575–578, 2000.

[42] A. Hamada, T. Furusawa, N. Saito, and E. Takeda, "A new aspect of mechanical stress effects in scaled MOS devices," *IEEE Trans. Electron Dev.*, pp. 895–900, 1991.

[43] H. S. Momose, T. Morimoto, K. Yamabe, and H. Iwai, "Relationship between mobility and residual-mechanical-stress as measured by Raman spectroscopy for intruded-oxide-gate MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 65–68, 1990.

[44] S. Tiwari, M. Fischetti, P. Mooney, and J. Welser, "Hole mobility improvement in silicon-on-insulator and bulk silicon transistors using local strain," *IEEE IEDM Tech. Dig.*, pp. 939–941, 1997.

[45] P. Kuo, J. L. Hoyt, and J. F. Gibbons, "Effects of strain on boron diffusion in Si and SiGe," *Appl. Phys. Lett.*, vol. 66, pp. 580–582, 1995.

[46] J. Kim, T. Kim, J. Park, W. Kim, B. Hong, and G. Yoon, "A shallow trench isolation using nitric oxide (NO)-annealed wall oxide to suppress inverse narrow width effect," *IEEE Electron Dev. Lett.*, vol. 21, pp. 575–577, 2000.

[47] T. Kuroi, T. Uchida, K. Horita, M. Sakai, Y. Itoh, Y. Inoue, and T. Nishimura, "Stress analysis of shallow trench isolation for 256 M DRAM and beyond," *IEEE IEDM Tech. Dig.*, pp. 141–144, 1998.

[48] M. Miyamoto, H. Ohta, Y. Kumagai, Y. Sonobe, K. Ishibashi, and Y. Tainaka, "Impact of reducing STI-induced stress on layout dependence of MOSFET characteristics," *IEEE Trans. Electron Dev.*, vol. 51, pp. 440–443, 2004.

[49] J. G. Fiorenza, G. Braithwaite, C. W. Leitz, M. T. Currie, J. Yap, F. Singaporewala, V. K. Yang, T. A. Langdo, J. Carlin, M. Somerville, A. Lochtefeld, H. Badawi, and M. T. Bulsara, "Film thickness constraints for manufacturable strained silicon CMOS," *Semicond. Sci. Technol.*, vol. 19, pp. L4–L8, 2004.

[50] T. Sato, Y. Takeishi, and H. Hara, "Mobility anisotropy of electrons in inversion layers on oxidized silicon surface," *Phys. Rev. B*, vol. 4, pp. 1951–1960, 1971.

[51] M. Kinugawa, M. Kakumu, and J. Matsunaga, "Submicron 3D surface-orientation-optimized CMOS technology," *IEEE VLSI Tech. Symp. Dig.*, pp. 17–18, 1986.

[52] S. Sugawa, I. Ohshima, H. Ishino, Y. Saito, M. Hirayama, and T. Ohmi, "Advantage of silicon nitride gate insulator transistor by using

microwave-excited high-density plasma for applying 100 nm technology node," *IEEE IEDM Tech. Dig.*, pp. 817–820, 2001.

[53] H. S. Momose, T. Ohguro, K. Kojima, S. Nakamura, and Y. Toyoshima, "110 GHz cutoff frequency of ultrathin gate oxide p-MOSFETs on (110) surface-oriented Si substrate," *IEEE VLSI Tech. Symp. Dig.*, pp. 156–157, 2002.

[54] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the universality of the inversion layer mobility in Si MOSFETs: Part I: Effects of surface orientation," *IEEE Trans. Electron Dev.*, vol. 41, pp. 2363–2368, 1994.

[55] T. Mizuno, N. Sugiyama, T. Tezuka, and S. Takagi, "[110] strained-SOI n-MOSFETs with higher electron mobility," *IEEE Electron Dev. Lett.*, vol. 24, pp. 266–268, 2003.

[56] T. Tezuka, N. Sugiyama, T. Mizuno, and S. Takagi, "High-performance strained Si-on-insulator MOSFETs by novel fabrication process utilizing Ge-condensation technique," *IEEE VLSI Tech. Symp. Dig.*, pp. 96–97, 2002.

[57] S. Takagi, A. Toriumi, M. Iwase, and H. Tango,, "On the universality of inversion layer mobility in Si MOSFETs: Part I - Effects of substrate impurity concentration," *IEEE Trans. Electron Dev.*, vol. 41, pp. 2357–2362, 1994.

[58] F. M. Bufler, A. Schenk, and W. Fichtner, "Strained-Si single-gate versus unstrained-Si double-gate MOSFETs," *Semicond. Sci. Technol.*, vol. 19, pp. S122–S124, 2004.

[59] L. Huang, J. O. Chu, S. A. Goma, C. P. D'Emic, S. J. Koester, D. F. Canaperi, P. M. Mooney, S. A. Cordes, J. L. Speidell, R. M. Anderson, and H.-S. P. Wong, "Electron and hole mobility enhancement in strained SOI by wafer bonding," *IEEE Trans. Electron Dev.*, vol. 49, pp. 1566–1571, 2002.

[60] L.-J. Huang, J. O. Chu, S. Goma, C. P. D'Emic, S. J. Koester, D. F. Canaperi, P. M. Mooney, S. A. Cordes, J. L. Speidell, R. M. Anderson, and H.-S. P. Wong, "Carrier mobility enhancement in strained Si-on-insulator fabricated by wafer bonding," *IEEE VLSI Tech. Symp. Dig.*, pp. 57–58, 2001.

[61] T. Mizuno, N. Sugiyama, T. Tezuka, and S. Takagi, "Novel SOI p-Channel MOSFETs with higher strain in Si channel using double SiGe heterostructures," *IEEE Trans. Electron Dev.*, vol. 49, pp. 7–14, 2002.

[62] T. Mizuno, S. Takagi, N. Sugiyama, H. Satake, A. Kurobe, and A. Toriumi, "Electron and hole mobility enhancement in strained-Si MOSFETs on SiGe-on-insulator substrates fabricated by SIMOX technology," *IEEE Electron Dev. Lett.*, vol. 21, pp. 230–232, 2000.

[63] Z.-Y. Cheng, M. T. Currie, C. W. Leitz, G. Tarashci, E. A. Fitzgerald, J. L. Hoyt, and D. A. Antoniadis, "Electron mobility enhancement in strained-Si n-MOSFETs fabricated on SiGe-on-insulator (SGOI) Substrates," *IEEE Electron Dev. Lett.*, vol. 22, pp. 321–323, 2001.

[64] Z. Cheng, J. Jung, M. L. Lee, H. Nayfeh, A. J. Pitera, J. L. Hoyt, E. A. Fitzgerald, and D. A. Antoniadis, "SiGe-on-insulator (SGOI): two structures for CMOS application," *SMA Symp.*, 2003.

[65] G. Hock, E. Kohn, C. Rosenblad, H. von Kanel, H.-J. Herzog, and U. Konig, "High hole mobility in $Si_{0.17}Ge_{0.83}$ channel metal-oxide-semiconductor field-effect transistors grown by plasma-enhanced chemical vapor deposition," *Appl. Phys. Lett.*, vol. 76, pp. 3920–3922, 2000.

[66] D. A. Antoniadis, I. Aberg, C. Ni Chleirigh, O. M. Nayfeh, A. Khakifirooz, and J. L. Hoyt, "Continuous MOSFET performance increase with device scaling: the role of strain and channel material innovations," *IBM J. Res. Develop.*, vol. 50, pp. 1–14, 2006.

[67] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, "Hole mobility enhancements in strained-Si/$Si_{1-y}Ge_y$ p-type metal-oxide-semiconductor field-effect transistors grown on relaxed $Si_{1-x}Ge_x$ (x < y) virtual substrates," *Appl. Phys. Lett.*, vol. 79, pp. 4246–4248, 2001.

[68] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors." San Jose, Calif., 2005.

[69] C. K. Maiti and G. A. Armstrong, *Applications of Silicon-Germanium Heterostructure Devices.* Institute of Physics Publishing, UK, 2001.

[70] W. Shockley, "Problems related to p-n junctions in silicon," *Solid-State Electron.*, vol. 2, pp. 35–60, 1961.

[71] S. M. Sze, *Physics of Semiconductor Devices.* John Wiley & Sons, New York, 2nd ed., 1981.

[72] T. Yaur and T. H. Ning, *Fundamentals of Modern VLSI Devices.* Cambridge University Press, 1998.

[73] J. Lee, A. L. Gutierrez-Aitken, S. H. Li, and P. K. Bhattacharya, "Responsivity and impact ionization coefficients of $Si_{1-x}Ge_x$ photodiodes," *IEEE Trans. Electron Dev.*, vol. 43, pp. 977–981, 1996.

[74] N. S. Waldron, A. J. Pitera, M. L. Lee, E. A. Fitzgerald, and J. A. del Alamo, "Impact ionization in strained-Si/SiGe heterostructures," *IEEE IEDM Tech. Dig.*, pp. 813–816, 2003.

[75] A. Neviani, G. Meneghesso, E. Zanoni, M. Hafizi, and C. Canali, "Positive temperature dependence of the electron impact ionization coefficient in $In_{0.53}Ga_{0.47}As$/InP HBTs," *IEEE Electron Dev. Lett.*, vol. 18, pp. 619–621, 1997.

[76] W. McMahon, A. Haggag, and K. Hess, "Reliability scaling issues for nanoscale devices," *IEEE Trans. Nanotech.*, vol. 2, pp. 33–38, 2003.

[77] E. Li, E. Rosenbaum, J. Tao, and P. Fang, "Projecting lifetime of deep submicron MOSFETs," *IEEE Trans. Electron Dev.*, vol. 48, pp. 671–678, 2001.

[78] K. A. Jenkins and K. Rim, "Measurement of the effect of self-heating in strained-silicon MOSFETs," *IEEE Electron Dev. Lett.*, vol. 23, pp. 360–362, 2002.

[79] K. A. Jenkins and J. Y. Sun, "Measurement of I-V curves of silicon-on-insulator (SOI) MOSFETs without self-heating," *IEEE Electron Dev. Lett.*, vol. 16, pp. 145–147, 1995.

# 8

## *Modeling and Simulation of Hetero-FETs*

Performance improvements for each new silicon CMOS technology generation have been driven mainly by geometric scaling, which implied increasing the channel doping concentration and vertical electrical field required to control short-channel effects. However, these two aspects of geometric scaling adversely affect carrier mobility due to impurity scattering and high-field saturation. Over the past decade, Technology CAD (TCAD) has become essential for developing and optimizing semiconductor process technologies ranging from nanoscale microprocessors to large-scale high-voltage power devices. TCAD is being increasingly used in manufacturing for advanced process control and parametric yield improvement. TCAD can lower technology development costs up to 40% by reducing the number of experimental lots and shortening development time. This is significant, considering the rising costs of product development and new wafer fabrication facilities. Process engineers are now struggling to integrate new materials and device structures into the process, often resulting in new physical effects such as process-induced stress, statistical dopant fluctuations, spatial quantization, and nonequilibrium transport. While in the past many of these effects could have been neglected, they are now primary considerations in developing a new process technology. With its physical basis, TCAD is capable of capturing detailed process effects, and calibrated TCAD flows can be used to generate computationally efficient process compact models (PCMs) that retain key process-to-device correlations. Using PCMs, manufacturing engineers can analyze process sensitivity, and identify key process steps to improve overall process capability.

The latest CMOS technology nodes rely heavily on stress engineering. From a general device performance perspective, stress impacts the carrier mobility through changes in the band structure and it also modulates the relative occupancy of the various valleys, as well as the effective masses and scattering rates. It is important to have accurate models that link the stress to the mobility enhancements. Several commercial simulators now offer a range of models that account for stress-induced changes to the carrier mobility. It includes a phenomenological model based on piezoresistivity and physics-based models. The underlying physics of the stress-induced mobility changes for holes is considerably more complicated than the physics for electrons and is not well captured by the aforementioned models.

Simulations of ever-shorter gate-length MOSFETs create more challenges

for using electronic transport models based on moments of the Boltzmann transport equation. These models assume quasiequilibrium and have limited ability to capture nonlocal effects. The full-band Monte Carlo method can describe highly nonlinear quasiballistic carrier transport. Full-band structures allow for the correct treatment of transport away from the edges of the energy bands and of high-energy effects such as impact ionization. In this chapter, we shall present a simulation approach to model various strained-Si MOSFETs on virtual substrates using available commercial simulator like, Silvaco. Process-induced stress which has been used to improve MOSFET performance will also be discussed. Using SENTAURUS, we discuss on a model-based understanding of the stress in device architectures, demonstrating how stress modeling can be used to understand stress effects in advanced CMOS strained-Si technologies.

## 8.1 Simulation of Hetero-FETs

It is well known that silicon-on-insulator devices provide benefits of reduced parasitic capacitance allowing for high-speed operation while minimizing power dissipation. Ultrathin body SOI devices have the added benefit of improved electrostatic integrity, and thus can be scaled to the shortest channel lengths. Enhancement of electron mobilities and hole mobilities on SOI are found to be $> 50\%$ and 15-20%, respectively [1]. However, although theoretical studies have predicted higher carrier mobility in ultrathin single-gate SOI samples compared to standard Si devices, this mobility increase (estimated in the best case at 10%) is obtained only for very low silicon layer thicknesses ($<5$ nm) and for a high inversion charge concentration. In other cases, i.e., a greater silicon thicknesses, a mobility degradation is experimentally observed [2].

One may wonder whether it would be possible to combine the two structures (strained-Si inversion layer and SOI inversion layers) to enjoy the advantages of each and at the same time, overcome the deficiencies they present separately. The advanced SOI MOSFETs with strained-Si channel on SiGe-on-insulator substrate (SSGOI MOSFETs) have been fabricated by combining the SIMOX (separation-by-implanted-oxygen) technology with regrowth of strained-Si films [3]. It is demonstrated that strained-SOI n- and p-channel MOSFETs have mobilities 1.6 and 1.3 times higher than conventional Si MOSFETs, respectively. Si/SiGe-on-insulator structures provide a good control of short-channel effects, have a lower parasitic capacitance and higher radiation tolerance and, moreover, present mobility values that are much higher than those found in conventional SOI MOSFETs. However, an accurate knowledge of high- and low-field transport in strained-Si for evaluating its true potential for submicron CMOS devices under the impact of nonequilibrium transport is

still lacking. In the following, simulation results of a comprehensive study on the transport properties of carriers in advanced SOI MOSFETs with strained-Si channel on SiGe-on-insulator substrate at low- and high-electric fields using a 2D device simulator (Silvaco-ATLAS) are presented.

## 8.2 Strained-Si Material Parameters for Modeling

Scaling of Si CMOS transistors to sub-50 nm gate length requires increased channel doping to control short-channel effects. Experimental measurements and theoretical predictions show degraded electron mobility due to high channel doping and is attributed to increased Coulomb interactions [4, 5]. The reduction in electron mobility caused by Coulomb scattering is so important that it is becoming one of the main limiting factors that have to be dealt with in order to continue the device downscaling.

At low fields, carrier mobility is dominated by Coulomb scattering, which is more effectively screened out at higher fields and thus, higher carrier densities in the inversion layer. Gamiz et al. [6] have shown that the contribution of the reduction of intervalley scattering is not as important as the contribution of the reduction of conductivity effective mass in the low longitudinal electric field transport regime, rather the reduction of the conduction effective mass is the main factor responsible for low-field mobility improvement in strained-Si. At moderate fields, the carrier mobility is determined by phonon scattering. Phonon scattering depends on $E_{eff}{}^{-0.3}$ for electrons, and has a slightly higher dependence for holes. Finally, in the high-field regime, surface roughness scattering dominates carrier mobility, with a dependence of $E_{eff}{}^{-2}$ for electrons and $E_{eff}{}^{-1}$ for holes.

Both theoretical and experimental studies have shown electron mobility and hole mobility enhancements when silicon is grown pseudomorphically on relaxed-$Si_{1-x}Ge_x$ [?]. In the conduction band (Fig. 8.1(a)), tensile strain splits the six-fold degeneracy, and lowers the two-fold degenerate perpendicular Δ-valleys with respect to the four-fold in-plane Δ-valleys in energy space. Two lowered valleys with the longitudinal effective mass axis perpendicular to the interface, and four raised valleys with the longitudinal mass axis parallel to the interface (see Fig. 8.1(a)) give rise to an enhancement in electron mobility in strained-silicon channel.

Such energy splitting suppresses intervalley carrier scattering between the two-fold and four-fold degenerate valleys, and causes preferential occupation of the two-fold valleys where the in-plane conduction mass is lower. Similarly, the effect of tensile strain on the valence band is shown in Fig. 8.1(b). The degeneracy between the heavy hole (HH) and light hole (LH) bands lifted, and valence band maximum is higher energy as compared to the unstrained

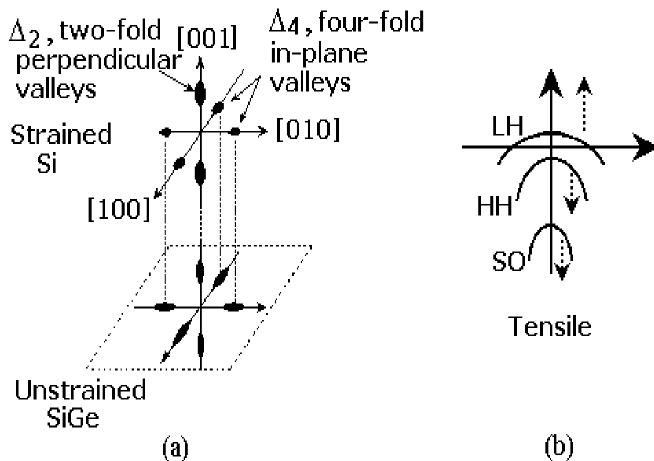



**FIGURE 8.1**
Schematic diagram of tensile strain-induced conduction and valence band splitting of strained-Si on $Si_{1-x}Ge_x$ substrate. After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

case. Band structure calculations using k.p method show that degeneracy is lifted at k = 0 and the curvature of the valence band increases, thus decreasing the effective mass near the Γ. The combination of a lower effective mass and reduced intervalley scattering gives rise to higher electron/hole mobility. Moreover, the lower intervalley scattering rates make energy-relaxation times higher, originating spectacular electron velocity overshoot [8]. Therefore, several material parameters for strained-Si and strained/relaxed-SiGe are needed for the device modeling viz., the band offsets encountered in strained layers, bandgap narrowing due to strain, doping-induced bandgap narrowing and the carrier mobility. Silicon strained in biaxial tension results in a type-II band offset and this allows tailoring of the band structure to confine both holes and electrons. Under biaxial strain, the strained-Si bandgap is reduced as follows [9]:

$$E_g(x) = 1.1 - 0.4x \tag{8.1}$$

where x is the Ge content in the top part of a completely relaxed-SiGe buffer.

To obtain accurate results for p-MOSFET simulations using DD model, it is necessary to account for effects associated with inversion layers. This is done by modeling mobility as having a dependence on the transverse and longitudinal fields. The reduction of channel mobility at high longitudinal-

electric fields has been taken into account via Eqn. 8.2.

$$\mu(E_{\parallel}) = \frac{\mu_0}{\left[1+\left(\frac{\mu_0 \times E_{\parallel}}{v_{sat}}\right)^{\beta}\right]^{1/\beta}} \tag{8.2}$$

where $\mu_0$ being the low-field mobility, $E_{\parallel}$ the longitudinal-electric field (along the channel), and $E_{\perp}$ the transverse-electric field (perpendicular to the silicon-$SiO_2$ interface). Note $v_{sat}$ and $\beta$ are fitting parameters, where $v_{sat}$ is the saturation velocity of the electrons in the channel. It has been shown that the value of $v_{sat} = 1.1\times10^7$ cm/s (T = 300 K) is the same regardless of the germanium mole fraction of the relaxed-SiGe alloy underneath the strained-Si layer.

To take into account the enhanced mobility in strained-Si with increasing strain (i.e., with Ge mole fraction, x, in the substrate) and the effects associated with inversion layers, the modified transverse field dependence CVT mobility model was taken in the simulation. The CVT model [10] is modified using an analytical expression involving Ge content, x as follows

$$\mu_{strained-Si}(E_{\perp}, T, N_A, x) = \mu_{CVT}(E_{\perp}, T, N_A)(1 + 4.31x - 2.28x^2) \tag{8.3}$$

The carrier mobility is approximated by the sum of three following terms,

$$\frac{1}{\mu_{CVT}(E_{\perp}, T, N_A)} = \frac{1}{\mu_{ac}} + \frac{1}{\mu_b} + \frac{1}{\mu_{sr}} \tag{8.4}$$

where $\mu_{ac}$ is the carrier mobility limited by scattering with surface acoustics phonons, $\mu_b$ is the carrier mobility in bulk silicon, and $\mu_{sr}$ is the carrier mobility limited by surface roughness scattering.

In the two-dimensional deformation potential theory of surface phonon scattering and for a nondegenerate surface, the electron mobility is given by:

$$\mu_{ac} = \frac{q\hbar^3 \rho u_l^2}{m^* m_{\mu} Z_A^2 k_B T} \tag{8.5}$$

where q is the elementary charge, $\hbar$ is the Dirac constant, $u_l$ is the sound velocity, $m^*$ and $m_{\mu}$ are the effective mass and the mobility mass, respectively, $Z_A$ is the deformation potential, $k_B$ is the Boltzmann's constant, and T is the absolute temperature. The carrier mobility limited by surface roughness scattering is given by:

$$\mu_{sr}(E_{\perp}, T) = \left(B\frac{T}{E_{\perp}} + C\frac{1}{E_{\perp}^{1/3}}\right)\frac{1}{T} \tag{8.6}$$

where $B_{init} = 3.1\times10^8$ cm/s and $C_{init} = 3.0\times10^7$ $(V/cm)^{-2/3}\cdot K\cdot cm/s$.
The carrier mobility in bulk silicon is given by:

$$\mu_b(N_A,T) = \left[\mu_o + \frac{\mu_{max}\left(\frac{T}{300}\right)^{-\gamma} - \mu_o}{1+(N_A/C_r)^{\alpha}} - \frac{\mu_1}{1+(C_s/N_A)^{\beta}}\right] \tag{8.7}$$

where $\mu_{max}$ are the expected maximum mobility values and $N_A$ is the local acceptor concentration. $\gamma_{init} = 2.42$ is an initial estimate for $\gamma$. If the trap density at the interfaces with the $SiO_2$ is kept low, Coulomb scattering is very weak, and its effects could be ignored. Therefore, only phonon scattering and surface-roughness scattering are taken into account the simulation. In addition, we have also ignored alloy scattering. Electrons/holes are also scattered by the random nature of the SiGe alloy. This is a fundamental limitation that cannot be removed (unless the alloy can be grown in an ordered form), but it is expected to have a weak effect in Si/SiGe system because only the tail of the electron distribution penetrates the SiGe layer [11].

## 8.3 Simulation of Strained-Si n-MOSFETs

Use of strained-Si in the channel of a MOSFET has emerged as one of the most promising approaches to enhance CMOS technology because of the improved transport properties and the compatibility with silicon technology [12, 13]. The strain dependence of electron inversion layer mobility has been measured in n-MOSFETs fabricated on a strained-Si/relaxed-$Si_{1-x}Ge_x$ heterostructure [14, 15]. For a Ge content of 20% in the underlying relaxed-$Si_{1-x}Ge_x$, the electron mobility the strained-Si is enhanced by roughly 75%. A hole mobility enhancement of similar magnitude has been reported for surface-channel strained-Si p-MOSFETs [16, 17]. These surface channel strained-Si n- and p-MOSFETs are identical to conventional Si MOSFETs in their operation principle, and are expected to retain all the advantages of surface-channel devices, including excellent current drive/turn-off characteristics, and scaling behavior.

n-MOSFETs fabricated using strained-Si deposited directly on insulator (SSDOI) are one of the most promising strained-Si device structures [18, 19] because they combine the advantages of both the strained-Si and SOI technology, avoid the problems of misfit dislocations associated with SiGe substrates, and do not require high channel doping for threshold voltage adjustment. In the following, the capability of the Monte Carlo device simulator SPARTA [20, 21, 22] to simulate devices under biaxial tensile strain is discussed. In addition, the use of parameter files for strained-Si, which contains transport parameters as a function of strain, is also illustrated. The device considered here is an advanced 25-nm strained-Si device fabricated directly on insulator

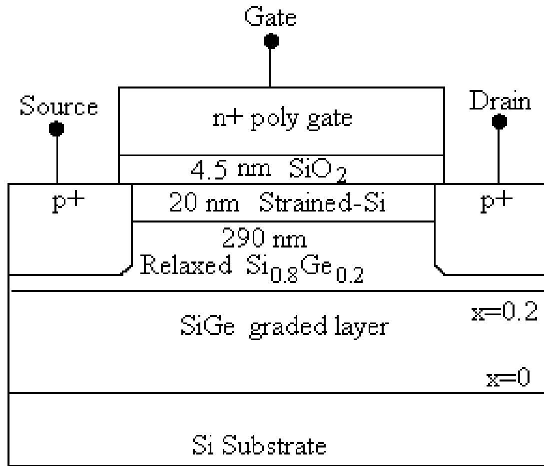



**FIGURE 8.2**
Schematic diagram of a MOSFET on pseudomorphic strained-Si grown on fully relaxed-SiGe buffer layer. After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

(SSDOI) n-MOSFET, which corresponds approximately to a scaled version of the realization of the SSDOI MOSFETs [18].

In simulation, Silvaco tools were used for process-to-device simulation. The process steps used in simulation were taken from the device fabrication steps reported in reference [23]. The strained-Si devices were fabricated using strained-Si epitaxial layers grown on relaxed-$Si_{0.8}Ge_{0.2}$ by gas source MBE at 800$^o$C. Fig. 8.2 shows the schematic cross-section of the strained-Si n-MOSFETs device structure. After epitaxial growth, MOSFETs were fabricated using a standard, self-aligned process. The channel was formed by boron implantation at 20 keV with a dose of $4\times10^{12}$/cm$^2$. An n$^+$-poly-gate process has been used to fabricate the devices, since it allows higher hole density in channel and increases the cross-over voltage [24] compared to conventional p$^+$-poly-gate devices. Gate spacer formation, source and drain (S/D) implantation, and a Ti salicide process were followed by metallization. The n$^+$ S/D regions were formed by As implantation at 30 keV with a dose of $3\times10^{15}$/cm$^2$. Strained-Si n-MOSFETs devices were fabricated with channel length (L) varying from 0.8 to 10 $\mu$m for channel width (W) from 5-100 $\mu$m.

Fig. 8.3 shows the output characteristics for a strained-Si n-MOSFETs device with x = 0.2, L = 10 $\mu$m, and W = 20 $\mu$m. The curves exhibit well-behaved long-channel MOSFET characteristics. There is very little output conductance, and the saturated drain current is increasing quadratically with the gate voltage, as expected for a long-channel MOSFET. The subthreshold

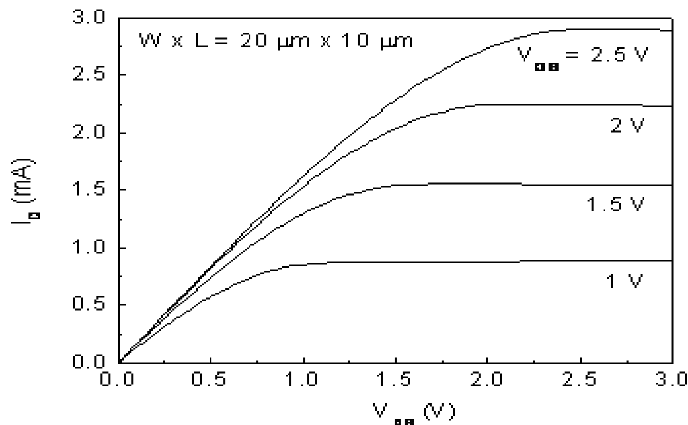



**FIGURE 8.3**
Output characteristics for n-MOSFETs with strained-Si on thick relaxed-$Si_{0.8}Ge_{0.2}$ buffer layer (W×L = 20 $\mu$m×10 $\mu$m). After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

characteristics of strained-Si n-MOSFETs devices with x = 0.2, L = 10 $\mu$m, and W = 20 $\mu$m is shown in Fig. 8.4 for four drain voltages, 0.1 V, 0.2 V, 1.5 V, and 2.5 V. The strained-Si MOSFETs show the normal subthreshold characteristics at the linear region. However, some amount of drain leakage currents is observed, and increases with the drain bias.

At lower gate voltage, the leakage current at the drain-source junctions of the strained-Si n-MOSFETs dominates the total drain current. The punch through effect is also observed at high drain bias, because of thicker strained-Si/SiGe layers. The thick SiGe buried layer, strained-Si channel and the Si setback layer can cause more source/drain punchthrough, and can contribute to a higher leakage current. This high leakage current is partly due to the high intrinsic carrier concentration of the SiGe layer. In addition, thermally induced degradation of the SiGe buried layer during device processing gives rise to threading dislocations which can enhance the leakage current. The subthreshold swing are 115 mV/decade for the strained-Si p-MOSFETs.

Fig. 8.5 shows the linear and saturation transconductances of the strained-Si n-MOSFETs as a function of gate voltage. For a given drain voltage the maximum transconductance $g_{msat}$ occurs at $V_{ds} = V_{Dsat} \equiv V_{gs}$-$V_t$, the onset of saturation, and is proportional to $V_{Dsat}$. The decrease in $g_m$ for higher gate voltages is largely due to source/drain resistance. The measured peak transconductance values are 7.5 and 70.0 mS/mm, at $V_{ds}$ = 0.1 V and 2.5 V, respectively.

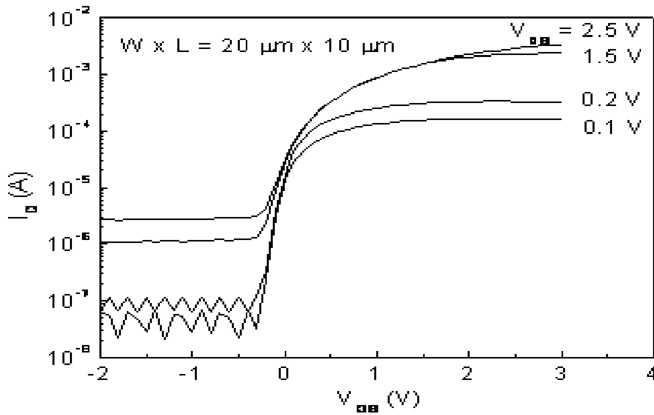



**FIGURE 8.4**
Subthreshold characteristics of strained-Si n-MOSFETs, where the drain bias are 0.1 V, 0.2 V, 1.5 V, and 2.5 V (W×L = 20 $\mu$m×10 $\mu$m.) After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

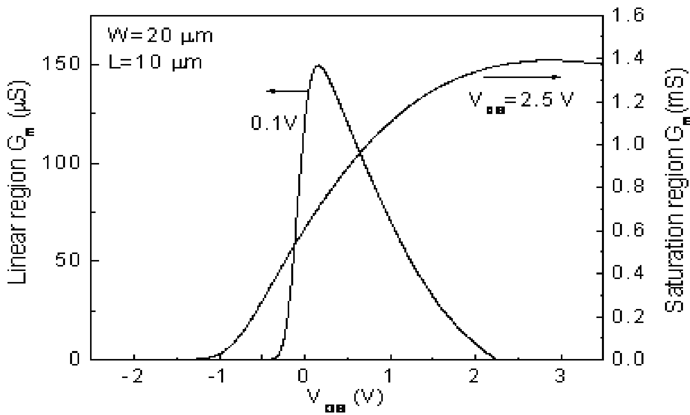



**FIGURE 8.5**
Linear and saturation transconductance of a W/L = 20 $\mu$m/10 $\mu$m strained-Si n-MOSFETs. After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

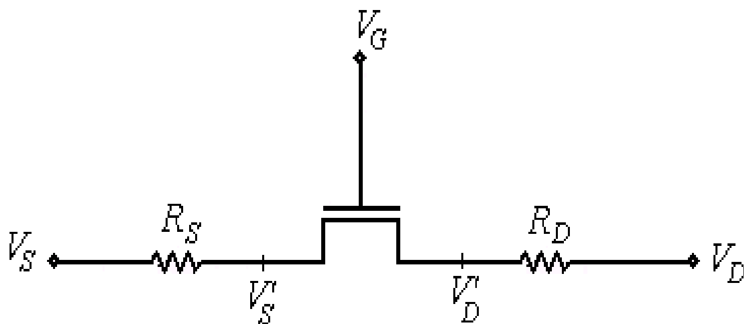



**FIGURE 8.6**
MOS transistor with source-and-drain resistance. After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

Source-drain junction regions of MOSFETs act as parasitic resistances limiting the performance of a device. In state-of-the-art IC technologies the parasitic source/drain series resistance ($R_{SD}$) of MOSFETs account for about 30% of reduction of the current drive capability (saturation current, $V_{gs}$ = $V_{supply}$ and $V_{ds}$ = 0.1 V). For the linear current (i.e., the drain current for $V_{gs}$ = $V_{supply}$ and $V_{ds}$ = 0.1V) the reduction due to $R_{SD}$ can be as high as 50% [25]. It is therefore important to study the source-drain resistances to get an accurate knowledge of the intrinsic performance of the silicon alloy devices.

Fig. 8.6 shows an MOS transistor with source-and-drain resistance $R_t$ = $R_s$ + $R_d$, where, $R_s$ is the source resistance and $R_d$ is the drain resistance. Effective drain and gate voltages, $V'_{ds}$ and $V'_{gs}$, respectively, are reduced below the voltages, $V_{ds}$ and $V_{gs}$ (applied at the external terminal of the device) due to the voltage drop across the parasitic resistors.

$$V'_{gs} = V_{gs} - I_{ds}R_s \tag{8.8}$$

$$V'_{ds} = V_{ds} - I_{ds}R_t \tag{8.9}$$

For a small $V_{ds}$, the drain current is represented in terms of the transconductance factor $\beta$ as [26]:

$$I_{ds} = \beta(V'_{gs} - V_t)V'_{ds} \tag{8.10}$$

where

$$\beta = \frac{\beta_0}{1 + U_0(V'_{gs} - V_t)} \tag{8.11}$$

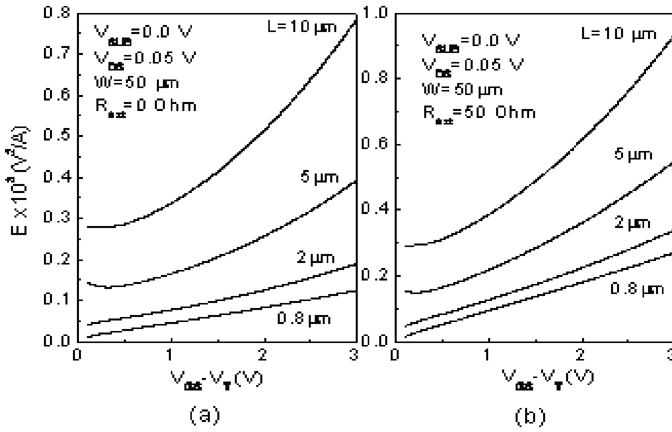



**FIGURE 8.7**
Variation of E vs. $V_{ds}$-$V_t$ for strained-Si n-MOSFETs with gate lengths of 0.8, 2.0, 5, and 10 $\mu$m at room temperature (when (a) no external and (b) 50 Ω external resistors were added in series with source). After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

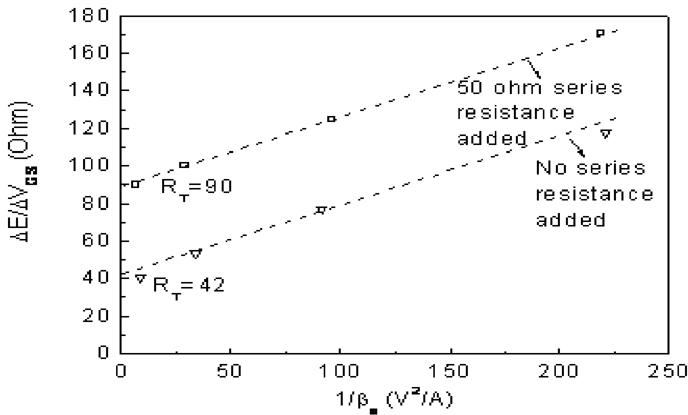



**FIGURE 8.8**
Slope $\Delta$E/$\Delta$$V_{gs}$ vs. $1/\beta_0$ for transistors with no external resistance and with 50 Ω added. After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

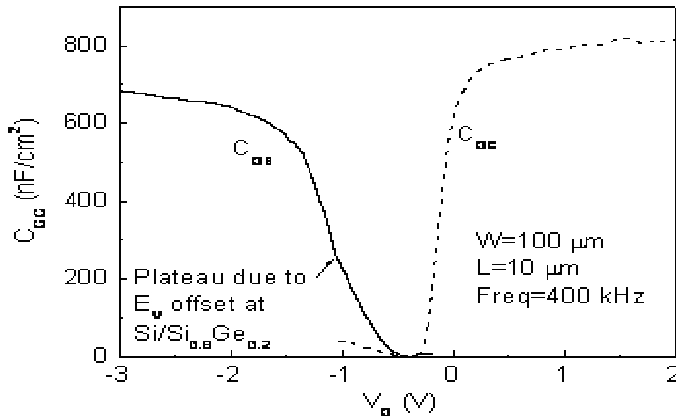



**FIGURE 8.9**
Split C-V for a large-area strained-Si n-MOSFETs. After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

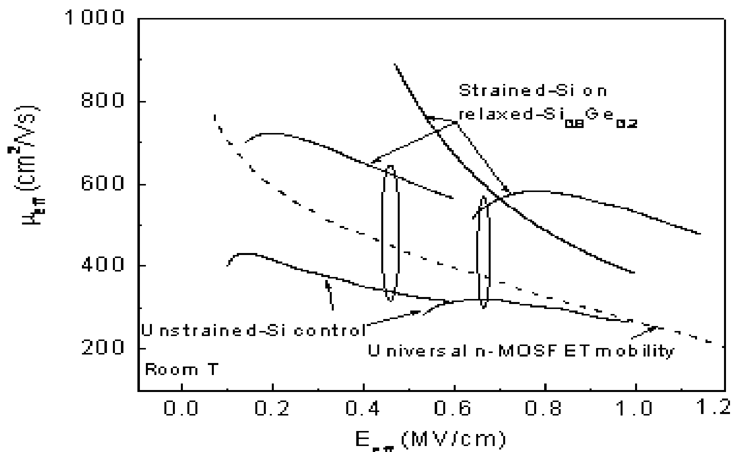



**FIGURE 8.10**
Effective mobility $\mu_{eff}$ as a function of vertical effective field $E_{eff}$. After S. K. Samanta, *Ultrathin Oxide and Oxynitride Gate Dielectric Films for Silicon Hetero-FETs*, Ph.D. Thesis, IIT Kharagpur, 2003.

$$\beta_0 = \frac{\mu_o C_{ox} W}{L} \tag{8.12}$$

Here, $\beta_0$ is the intrinsic transconductance, $U_0$ is the mobility degradation factor, $\mu_0$ is the low-field mobility, and $C_{ox}$ is the gate oxide capacitance per unit area. The threshold voltage ($V_t$) is commonly extracted by the usual method of extrapolating the linear portion of the plot of $I_{ds}$-$V_{gs}$ or $I_{ds}/\sqrt{g_m}$ vs. $V_{gs}$ at $V_{ds}$ = 0.1 V. Substituting the above relationships into Eqn. (8.10), we obtain a second-order equation in $I_{ds}$ with $V_{gs}$ and $V_{ds}$ as coefficients:

$$I_{ds}^2 R_s(U_0 + \beta_0 R_t) - I_{ds}[(U_0 + \beta_0 R_t)(V_{gs} - V_t) + 1 + \beta_0 V_{ds} R_s] + \beta_0 V_{ds}(V_{gs} - V_t) = 0 \tag{8.13}$$

Assuming $R_s V_{ds} \ll R_t(V_{gs}-V_t)$, and further simplifying Eqn. (8.13), it becomes [26]:

$$I_{ds} = \frac{\beta_0(V_{gs} - V_t)}{1 + (U_0 + \beta_0 R_t)(V_{gs} - V_t)} V_{ds} \tag{8.14}$$

Then, Eqn. (8.14) can be written as:

$$E = \frac{(V_{gs} - V_t)V_{ds}}{I_{ds}} = \frac{1 + (U_0 + \beta_0 R_t)(V_{gs} - V_t)}{\beta_0} \tag{8.15}$$

Expression E, at $V_{ds}$ = 0.1 V, is plotted as a function of $V_{gs}$-$V_t$ for strained-Si n-MOSFETs devices in Fig. 8.7(a). It is noted that E varies linearly with $V_{gs}$-$V_t$, as noted from Eqn. 8.15. Eqn. 8.15 based on the above assumption is not valid for $V_{ds}$-$V_t$, which is clear from the nonlinear behavior of the curves near $V_{ds}$-$V_t$ = 0. The intersections of the extrapolated straight lines with vertical axis (E) gives the value of $1/\beta_0$.

The slope E can then be written as:

$$\frac{dE}{dV_{gs}} = \frac{A}{\beta_0} = R_t + \frac{U_0}{\beta_0} \tag{8.16}$$

The slope of the curve ($dE/dV_{gs}$) from Fig. 8.7(a) is plotted with $1/\beta_0$ for the devices in Fig. 8.8. The slope of the resulting straight line yields a mobility degradation factor $U_0 = 0.4\ V^{-1}$ and the intercept yields a source-and-drain resistance $R_t = 42\ \Omega$.

Adding 50 Ω external resistance in series with the source of each transistor, E, at $V_{ds}$ = 0.1 V, is plotted as a function of $V_{gs}$-$V_t$ for strained-Si n-MOSFETs devices in Fig. 8.7(b). Similarly, the value of $dE/dV_{gs}$ as function of $1/\beta_0$ for the devices is plotted in Fig. 8.8. As seen in Fig. 8.8, the intersection of the line with vertical axis is at $R_t = 90\ \Omega$, close enough to the expected 92 Ω. The value of $U_0$ is unchanged. The higher $U_o$ values devices is due to higher surface roughness strained-Si [17]. However, the roughness of the strained-Si depends on the quality of the SiGe buffer and the Ge concentration.

The split C-V characteristics of a long channel (W×L = 100×10 $\mu m^2$) strained-Si device are shown in Fig. 8.9. The small plateau observed in the gate-to-substrate capacitance ($C_{GB}$) curve is caused by the valence band discontinuity at the strained-Si/relaxed-$Si_{1-x}Ge_x$ interface, which signifies hole confinement in a well formed by the potential barrier. The gate oxide thickness $t_{ox}$ determined from the maximum inversion capacitance was 45 Å for strained-Si devices.

The effective mobility $\mu_{eff}$ is derived from DC drain conductance measurements as reported by Sun and Plummer [27] using the relation:

$$\mu_{eff} = \frac{L}{W}\frac{G_d}{Q_{inv}} \tag{8.17}$$

where $g_d$ is the drain conductance measured on large devices. Low $V_{ds}$ (typically 10 mV) is used to minimize the effect of lateral field and nonuniform vertical field and carrier concentration along the channel. The effective field was obtained by

$$E_{eff} = \frac{1}{\epsilon_s}(Q_b + \eta Q_{inv}) \tag{8.18}$$

where $\eta = 1/2$ for electrons and 1/3 for hole mobility [27]. The inversion charge $Q_{inv}$ was found by integrating the $C_{GC}$ curve, measured by split C-V, using the following relation:

$$Q_{inv} = \int_{-\infty}^{V_{gs}} C_{GC} dV_{gs} \tag{8.19}$$

where $Q_b$ is the integrated maximum depletion charge.

Fig. 8.10 shows the comparison of effective mobility. The universal mobility curve for state-of-the-art Si n-MOSFETs [28] is shown in the figure (dotted line). The mobility data measured lightly doped ($2\times10^{16}$ $cm^{-3}$) strained-Si devices [14] and highly doped ($\sim \times10^{18}$ $cm^{-3}$) [29] are also shown for comparison. Through the vertical $E_{eff}$ range, the measured mobility of the strained-Si device exhibits a significant enhancement over that of epi Si-control wafer and the state-of-the art MOSFET's universal mobility. At 1 MV/cm, the strained-Si devices exhibit a 44% higher mobility than the universal mobility and the unstrained-Si control device. The effective mobility of the strained device exhibits an enhancement of roughly 80% over that of the control device at 0.6 MV/cm effective field.

The electron $\mu_{eff}$ enhancement observed at low lateral $E_{eff}$ for lightly doped devices appears to be sustained at higher $E_{eff}$ of > 0.5 MV/cm. This is consistent with the theoretical calculations which predicted enhancement of phonon-limited mobility in strained-Si even at high vertical field [28]. The enhancement factor of roughly 80% is also in a good agreement with the theoretical predictions of phonon-limited mobility [28]. The conduction band splitting between the two-fold and four-fold degenerate valleys reduces the

number of available final states for f-optical phonon scattering, suppressing the overall intervalley scattering. However, suppressed occupation of four-fold degenerate valleys, which have stronger interaction with intervalley phonons, contributes to the mobility enhancement.

## 8.4 Characterization of Strained-Si Hetero-FETs

The self-heating effect (SHE) characterization of MOSFETs and the accurate extraction of the equivalent thermal impedance of the device (thermal resistance, RTH and capacitance, CTH) are needed for advanced device and IC simulation. There is a lack of simple yet efficient extraction procedures for the evaluation of thermal parasitic components needed for advanced MOSFET simulation. One of the key phenomenon related to the operation of strained-Si MOSFETs are self-heating effect which is directly dependent on the power dissipated by the device. SHE results in a reduction of the drain current and the well-known negative output conductance effect. SHE characterization and its modeling originate from the SOI MOS devices where this effect is particularly critical [30, 31].

As SiGe has a lower thermal conductivity than Si [32], this causes heat to accumulate in heterostructure devices involving SiGe [33]. As the biaxially strained-Si devices are fabricated on SiGe virtual substrates, device self-heating occurs in the channel region of strained-Si/SiGe MOSFETs [34]. It has been reported by several workers that the self-heating lowers the performance of the devices in terms of fall in the drain current due to the accumulation of heat in the channel region [35].

Although much work on self-heating effects in SOI devices has been reported in the literature, so far very little information is available for strained-Si/SiGe MOSFETs. It is important to model the drain current for the strained-Si/SiGe device including the effects of self-heating. In the following, some aspects of analytical model development for strained-Si/SiGe MOSFETs including self-heating is discussed. In the following, we discuss the impact of self-heating effect (SHE) pulsed-measurement parameters (pulse width and duty factor) and propose an extraction method for the thermal parasitic network of the strained-Si devices.

Various pulsed measurements aimed to MOSFET characteristics without SHE have been reported. In general, such measurement setup requires a dedicated and/or expensive configuration since they cannot be achieved with semiconductor parameter analyzers where integration time exceeds 80 $\mu$s and much faster pulses are required to avoid SHE. It has been demonstrated that SHE is canceled by using pulses with duration less that 2 $\mu$s and duty factor lower that 1:100. The pulse-extraction method of device thermal resistance

and capacitance exploits analytical modeling and dedicated extraction plots using the measurements of output characteristics at various applied pulses and the gradual reduction of SHE with pulse duration and duty factor. Both RTH and CTH are assumed and extracted in saturation region as quasi-independent functions of the device bias at a given external temperature. A pulse generator is used to turn on and off the transistor by applying a square wave signal on the gate. The drain is biased through a fixed resistance by a high-voltage generator. A digital oscilloscope that triggers the signal applied on the gate is used for monitoring the voltage variations on the high-voltage generator and the drain of the transistor. A supply line capacitance is commonly used to reduce the peaks and stabilize at constant level the $V_{dd}$ voltage when transitions occur.

Jankovic et al. [36] have modeled the self-heating effect in strained-Si MOSFETs. The bulk-Si NQS MOS device model was successfully modified and implemented for modeling. The simulated and the measured transfer and output characteristics of otherwise identical SS n-MOSFETs on 15% Ge virtual substrate and bulk-Si device with different effective channel lengths are shown in Fig. 8.11. The excellent agreement is obtained with using a unified set of NQS MOS model parameters for modeling each of the experimental devices. From Fig. 8.11(b), it is seen that the negative output conductance is high visual in short-channel n-MOSFETs, whereas the same is not observed in the control-Si transistors.

## 8.5 TCAD: Strain-engineered Hetero-FETs

Strain engineering has become a key component of emerging technologies, however, due to the great complexity and cost in IC fabrication, it is difficult to physically develop new processes. Technology Computer Aided Design (TCAD) tools are invaluable for shortening new technology development and for optimizing existing processes [37]. Stress engineering has become the major approach besides conventional scaling to increase CMOS performance. Already at the 90-nm node, stress engineering has become indispensable to meet the performance targets of the International Technology Roadmap for Semiconductors (ITRS). The 2005 International Technology Roadmap for Semiconductors [38] states that the use of TCAD will provide as much as a 40% reduction in technology development costs. In this highly competitive environment, predictive power of TCAD to reduce cost and time in product development is highly useful. The knowledge gained through TCAD simulation can be used to optimize processes and to identify potential yield problems that need to be addressed in the design before volume production. TCAD is being used more and more efficiently to explore new device architectures

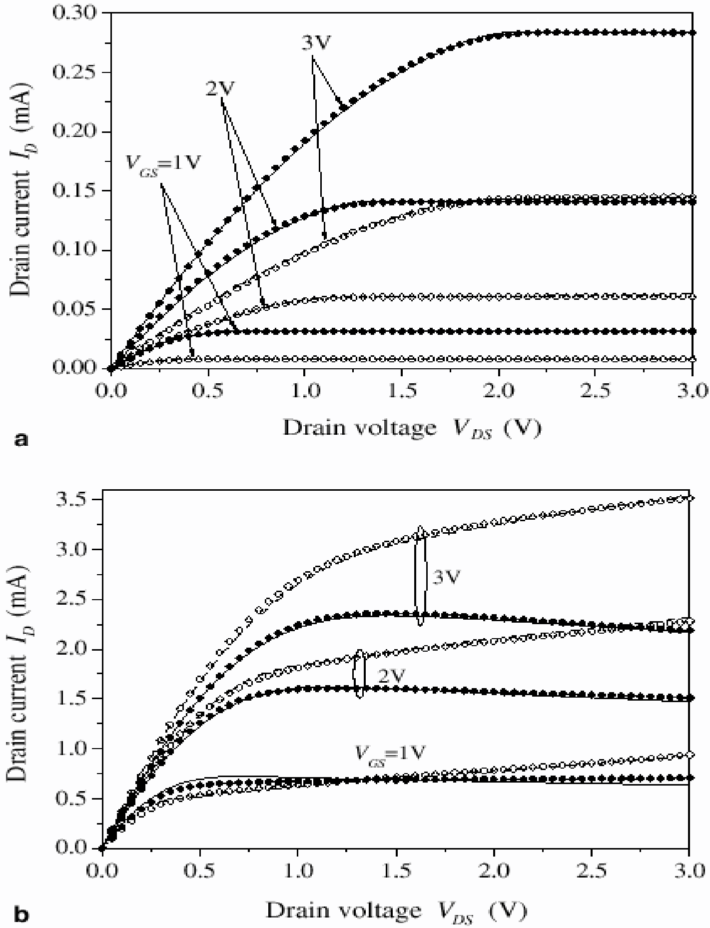



**FIGURE 8.11**

Comparisons of measured and modeled output characteristics of bulk-Si and SS MOSFETs with (a) $L_{eff}$ = 9.9 $\mu$m and (b) $L_{eff}$ = 0.1 $\mu$m. After N. D. Jankovic et al., *Solid-State Electron.*, Vol. 50, 2006(496-499). With permission.

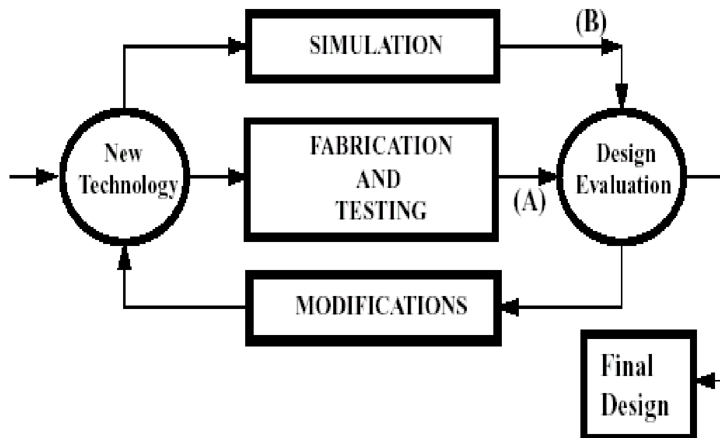



**FIGURE 8.12**
Simulation and experiment as alternative routes in development of new technologies. Route B would eventually involve a run through Route A, but helps in considerably lowering the time and expense of Route A. After S. Chaudhry, *Analysis and Modeling of Stress Related Effects in Scaled Silicon Technology*, Ph.D. Thesis, University of Florida, 1996.

and optimize process flows in manufacturing for advanced process control and parametric yield improvement.

In this section, we shall present a simulation approach to model front-end process-induced stress which has been used to improve device performance. The main sources of stress in advanced Si CMOS technologies and how they can be used to improve device performance are discussed in detail. We discuss a model-based understanding of the stress in device architectures, demonstrating how stress modeling can be used to understand stress effects in advanced CMOS strained-Si technologies.

Fig. 8.13 schematically illustrates the various software aids for IC technology development. Process simulation deals with all aspects of IC fabrication. With the proper input parameters (process recipes, layout geometries), process simulation determines the details of the resulting device structure, such as the boundaries of different materials and the distribution of impurity ions. The application of TCAD tools in the development of new processes and novel device structures involving strain-engineering has become challenging, an alternative to the experimental route (see path B in Fig. 8.12).

With scaled feature dimensions in the nanometer range, parametric variations become a dominant yield loss component. The transition to 90-nm technology proved to be extremely challenging, and it is expected to worsen at 65 nm and 45 nm. Calibrated TCAD flows can be used to generate compu-

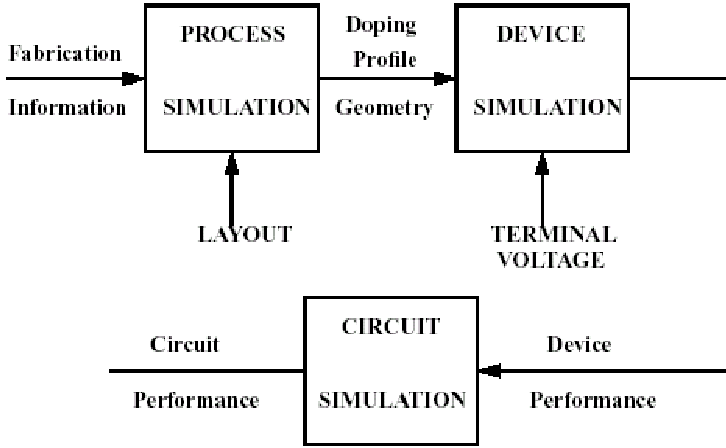



**FIGURE 8.13**
Process simulation as part of the general simulation methodology. After S. Chaudhry, *Analysis and Modeling of Stress Related Effects in Scaled Silicon Technology*, Ph.D. Thesis, University of Florida, 1996.

tationally efficient process compact models (PCMs) that retain key process-to-device correlations. Using PCMs, manufacturing engineers can analyze process sensitivity and identify key process steps to improve overall process capability.

Rapid advances in process technologies have created the need to understand new physical phenomena in order to optimize process structures and enhance device performance. For example, stress simulations have become indispensable in studying the effects of adding stress in the channel to enhance drive current. Three-dimensional simulation is needed to comprehend fully the impact of new physical phenomena on device performance.

Stress engineering poses significant new challenges for device simulation since important intrinsic material parameters such as mobility, bandgap, or quantization phenomena are changed under strain. At the same time, scaling enhances the influence of quasiballistic transport in the on-state, which is not captured by classical device simulation, but requires Monte Carlo simulation. As discussed in Chapter 4, from a device performance perspective, stress impacts the carrier mobility through changes in the band structure, and it also modulates the relative occupancy of the various valleys, as well as the effective masses and scattering rates. It is important to have accurate models that link the stress to the mobility enhancements. Sentaurus Device offers a range of models that account for stress-induced changes to the carrier mobility. It includes a phenomenological model based on piezoresistivity and physics-based models [37]. The underlying physics of the stress-induced mobility changes

for holes is considerably more complicated than the physics for electrons and is not well captured by the aforementioned models. In heterostructure devices and modern deep-submicron CMOS devices, the tunneling of carriers through barriers must be modeled accurately to obtain reliable predictions of device performance. The direct tunneling model is fully quantum-mechanical and can be applied to problems with simple barriers, such as gate leakage through a single-layer oxide. The most advanced tunneling model in Sentaurus Device can handle arbitrary barriers and is suitable for the simulation of gate leakage through multilayered gate dielectric stacks or tunneling through the barrier of a Schottky contact.

Simulations of ever-shorter gate-length MOSFETs create more challenges for using electronic transport models based on moments of the Boltzmann transport equation. These models assume quasiequilibrium and have limited ability to capture nonlocal effects. The full-band Monte Carlo method can describe highly nonlinear quasiballistic carrier transport. Full-band structures allow for the correct treatment of transport away from the edges of the energy bands and of high-energy effects such as impact ionization. The ensemble approach allows users to model transient and steady-state behavior and assess the deleterious effects of plasmons on saturation currents below 40 nm [5].

In a MOSFET, surface roughness scattering occurs in addition to phonon and impurity scattering. In SPARTA [39], this mechanism is modeled by a combination of specular and diffusive scattering where the percentage of diffusive scattering is 15%. Scattering is determined by the conservation of energy and parallel wave vector. The effect of strain enters through the modified band structure that affects the selection of the after-scattering wave vector when considering the above conservation laws. Fig. 8.14 shows the SPARTA results for the strain-induced improvement of the long-channel effective electron and hole mobility as a function of the effective field for several substrate germanium contents in comparison with measurements of MIT. The effective mobilities were extracted in the linear current regime of a 0.5 $\mu$m MOSFET from the Monte Carlo drain current similar to the procedure used in the measurement evaluation.

In the following, we present a methodology addressing these issues based on TCAD developments reported by W. Fichtner of SYNOPSYS [40]. Stress simulation for mobility enhancement involves the following:

- Simulation of fabrication process
  - Device structure generation via etching and deposition steps
- Modeling stress generation and evolution
  - Unintentional stress
    - Intrinsic, thermal mismatch, expansion and shrinkage, relaxation
  - Intentional stress
    - SiGe lattice mismatch, cap layer intrinsic
- Modeling dopant diffusion under stress
- Modeling of device characteristics
  - Simulated structure with dopant profiles and stress distributions

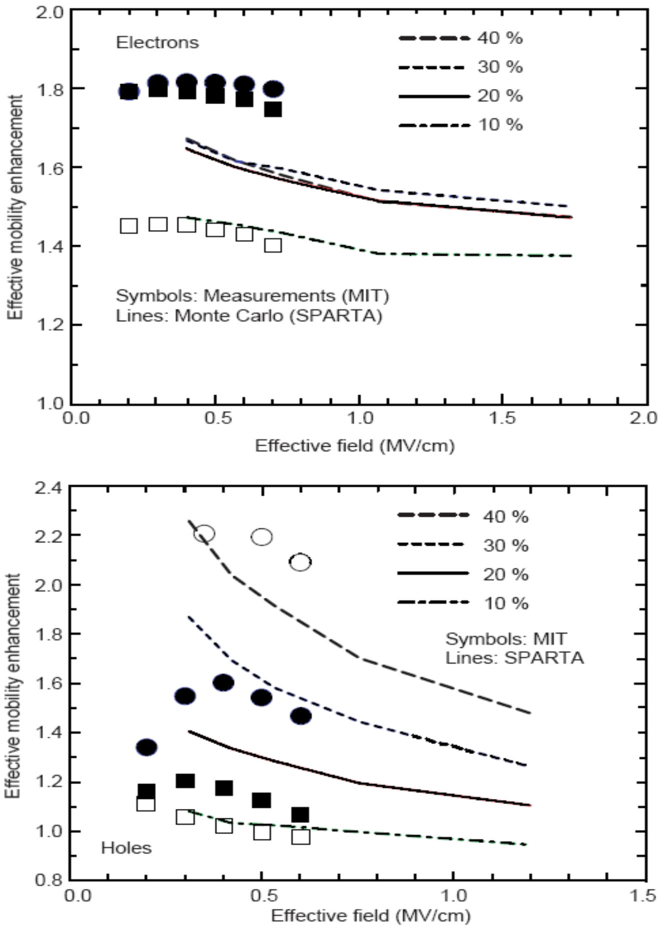



**FIGURE 8.14**

Enhancement of electron (top) and hole (bottom) long-channel effective mobility by biaxial tensile strain caused by growing a thin silicon film on a relaxed-$Si_{1-x}Ge_x$ substrate. Comparison is made between measurements by MIT and results of the Monte Carlo device simulator SPARTA. Source: TCAD News, Synopsys, August 2005 issue.

**TABLE 8.1**
Impact of stress on mobility computed using piezoresistive model.
Source: W. Fichtner, 2005 SiNANO Lecture Materials.

| Stress component | Tensile nMOS | Tensile pMOS | Compressive nMOS | Compressive pMOS |
|---|---|---|---|---|
| 1 GPa along channel (x) | +30% | -70% | -30% | +70% |
| 1 GPa along channel (Z) | +20% | +70% | -20% | -70% |
| 1 GPa Vertical (Y) | -50% | +1% | +50% | -1% |

Calibrated device model parameters
Validation, prediction and feedback for process optimization

Carrier mobility may be enhanced through process-induced channel stress engineering in the three following ways:

- Stress transfer depends on distance, geometry, process, and material
- Channel stress can be induced by multiple sources in a transistor
- Process stress simulator is a very effective tool for mobility engineering

While considering simulation of stress/strain effect in hetero-FETs, one needs to consider the following:

- Stress simulation for mobility engineering
  - Mechanical stress states
  - Stress impact on mobility
  - Stress-mobility model validation
- Stress assisted mobility engineering applications
  - Vertical stress engineering with cap layer
- Cap layer thickness
- Cap layer lamination
- Layout variation
- Spacer effect
- Recess effect
  - Lateral stress engineering with SID elevation
  - Packaging stress impact

Stress simulation involves several factors as listed below:

- Stress tensor $\sigma_{ij}$ has 6 components
  - Three normal components $\sigma_{xx}$, $\sigma_{yy}$, and $\sigma_{zz}$
  - Three shear components $\sigma_{xy}$, $\sigma_{xz}$, and $\sigma_{yz}$
- Normal stress states in channel center
  - Uniaxial stress state
    - One normal component dominant
  - Biaxial stress state
    - Two normal components dominant
  - Triaxial stress state
    - Three normal components dominant
- Sign conventions
  - Positive for tensile, negative for compressive

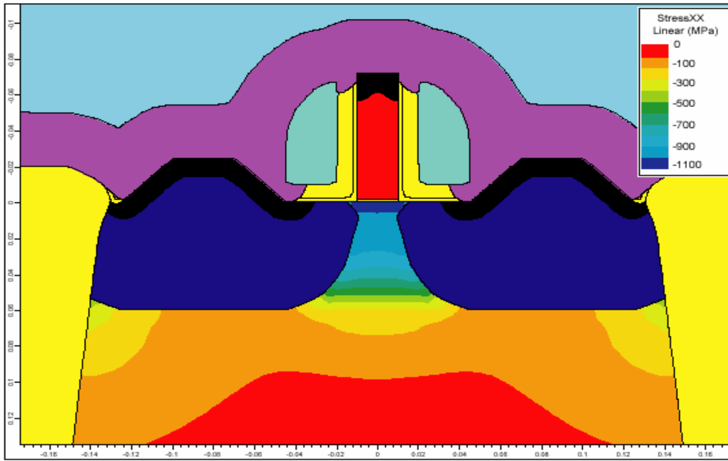



**FIGURE 8.15**
Compressive lateral stress in 20-nm pMOS with elevated SiGe S/D, and compressive spacer and cap layer. Stress is only shown in silicon. Source: TCAD News, Synopsys, August 2005.

TCAD stress engineering for nMOS and pMOS steps are summarized below:

- nMOS:
  - Most desirable channel stress
    - Compressive stress in the vertical direction
  - Most effective stress source
    - Tensile stretched cap layer
- pMOS:
  - Most desirable channel stress
    - Compressive stress along the current direction
  - Most effective stress source
    - Compressed SiGe lattice in S/D

The most beneficial stress components are different for nMOS and pMOS transistors. The nMOS benefits most from compressive vertical stress and also somewhat from tensile stress in the lateral direction, whereas the pMOS benefits most from the compressive stress in the direction along the channel. Tensile cap layer simultaneously introduces compressive vertical and tensile lateral stresses that enhance electron mobility. The rigid polysilicon gate readily transfers stress from the cap layer into the channel, while the softer metal gate transfers only half of the stress into the channel. The absence of the rigid nitride spacers helps to improve stress transfer efficiency through the poly-gate. The relatively soft metal gate adsorbs all stress and transfers nothing into the channel.

For a 20-nm p-MOSFET, it is possible to create lateral compressive stress

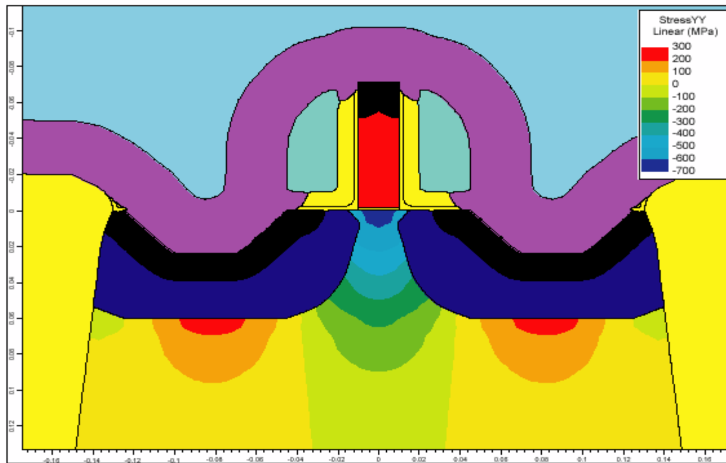



**FIGURE 8.16**
Compressive vertical stress in 20-nm nMOS with recessed SiC S/D, tensile silicide, tensile spacer, and tensile cap layer. Stress is only shown in silicon. Source: TCAD News, Synopsys, August 2005.

along the channel well above 1 GPa by using a combination of SiGe source and drain, elevated S/D, and compressive cap layer, as shown in Fig. 8.15. For this particular case, the low-field hole mobility increases by 91%, leading to $I_{on}$ enhancement of about 45%. Elevation of the SiGe S/D is instrumental in increasing the useful lateral compressive stress generated by the Si/SiGe lattice mismatch and simultaneously suppressing the harmful lateral stress generated by the tensile silicide.

Fig. 8.16 illustrates that for a 20-nm n-MOSFET, it is possible to create vertical compressive stress in the channel of over 600 MPa, plus about 400 MPa tensile stress along the channel by combining tensile cap layer, SiC S/D, recessed S/D, and tensile silicide. For this particular case, the low-field electron mobility increases by 44%, leading to $I_{on}$ enhancement of about 22%. Recessed S/D helps to increase the useful vertical compressive stress induced by the tensile cap layer, but simultaneously it reduces useful lateral tensile stress generated by the combination of SiC S/D and tensile silicide. Having a flat S/D would improve the lateral stress at the expense of the vertical stress component. Trade-offs like this are typical in stress engineering and require analysis of specific device geometry with specific stress sources to optimize the device performance.

## 8.6 SPICE Parameter Extraction

Hot-carrier degradation due to self-heating of strained-Si MOSFETs is an important reliability issue in deep submicron technology. Scaling and high-performance advantages make the device channel length shrink continuously. It is known that with the operating voltage downscaling, devices will be in the stressed condition. It has been reported that short-channel n-MOSFETs incorporating thin strained-Si (on relaxed-SiGe) surface channels can achieve significant drive current enhancement. Thus, strained-Si structures are promising candidates for enhancement of CMOS performance. However, so far little research has been done on strained-Si SPICE parameter extraction and reliability issues.

Currently, strained-Si research is mainly focused on device fabrication and technology development. However, as with all CMOS generations, there will eventually be strong interest in circuit applications. At this time, little is known about the SPICE parameter of strained-Si MOSFETs. Circuit simulation has become more and more important for the development of integrated circuits. SPICE modeling of MOSFETs is an essential aspect of circuit simulation. Usually, the circuit designer starts the design by using the SPICE parameters for a device. Threshold voltage, inversion layer carrier mobility, source and drain series resistance, together with device dimensions, form the major parameters of the SPICE-based submicrometer MOSFET circuit model used in the circuit simulation of MOS integrated circuits.

In the following, we discuss the thermal modeling for self-heating of strained-Si transistors using finite element simulation (ANSYS). A simple thermal model is used for parametric study of the self-heating in strained-Si device and the influence of the maximum temperature rise in the device, which is evident from the rise in temperature in the device, obtained using ANSYS thermal simulation, as shown in Fig. 8.17. We describe how one can extract the SPICE parameters of strained-Si MOSFETs both at room and low temperature using ICCAP tool. The extraction of threshold voltage, source-drain resistance, and gate field mobility reduction factor will be discussed. The experimental output characteristics of a strained-Si n-MOSFET is shown in Fig. 8.18. The SPICE parameters were extracted from Fig. 8.19 using Silvaco UT-MOST (BSIM3) tool. Several important SPICE parameters extracted from Fig. 8.19 are shown in Table 8.2.

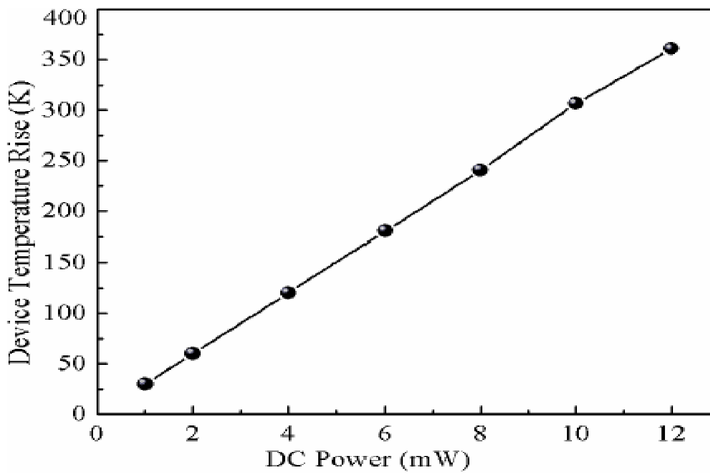



**FIGURE 8.17**
ANSYS-simulated temperature profile in a strained-Si device. After T. K. Maiti, unpublished data.

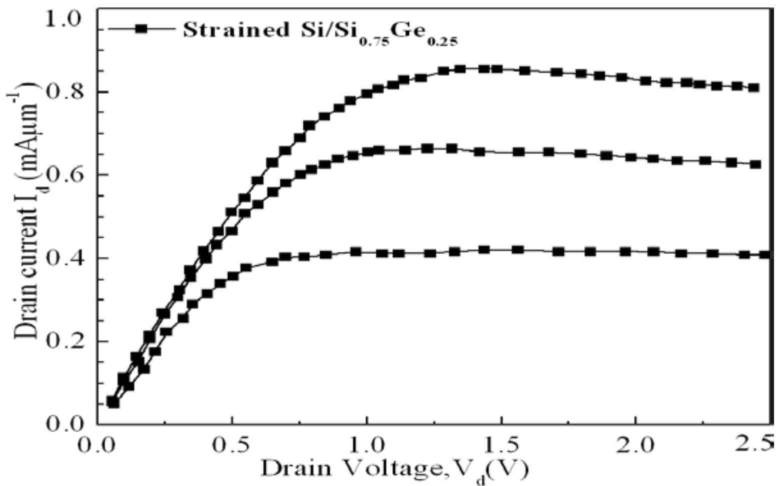



**FIGURE 8.18**
Experimental $I_d$-$V_d$ characteristics for $V_g$-$V_t$ = 1 V, 2 V, 3 V, respectively. After S. H. Olsen et al., *IEEE Trans. Electron Dev.*, Vol. 51, 2004(1156-1163). With permission.

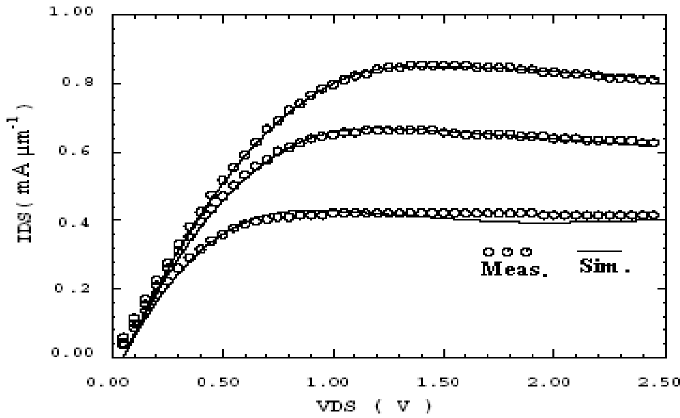



**FIGURE 8.19**
SmartSpice simulated output characteristics of a strained-Si n-MOSFET for parameter extraction. After S. S. Mahato, unpublished data.

**TABLE 8.2**
Several SPICE parameters for strained-Si n-MOSFETs extracted using ICCAP tool. After S. S. Mahato, unpublished data.

| Parameter | Value |
|---|---|
| VTH0(V) | 0.2000 |
| K1($V^{0.5}$) | 0.0500 |
| K2(-) | -0.0186 |
| K3(-) | 80.000 |
| W0(m) | 2.50E-6 |
| NLX(m) | 1.74E-7 |
| DVT2W($V^{-1}$) | -0.0320 |
| DVT0(-) | 8.200 |
| DVT1(-) | 0.2500 |
| U0($cm^2$/V-s) | 970.00 |
| KETA($V^{-1}$) | -0.0470 |
| RD(ohm/square) | 45.000 |
| PRWD1($V^{-1}$) | 10.000 |
| PRWD2($V^{-2}$) | 10.000 |
| UD(-) | -3.500 |
| VSATG(0+0.05$V^{-1}$) | -0.1000 |

## 8.7 Performance Assessment

Strained-Si channel devices have recently become of interest for future high-performance applications due to higher carrier mobility and preservation of conventional device structure/geometry. After successful fabrication of the sub-100 nm n- and p-type strained-Si MOSFETs, it is now possible to investigate the scalability of these devices. The device/circuit performance of strained-Si MOSFETs including strained-Si channel-on-insulator (SSOI) may be predicted via a physics-based compact model. Kim et al. [41, 42, 43] have assessed the circuit performance of strained-Si devices including SSOI via a physics-based circuit model calibrated against fabricated 70-nm strained-Si and unstrained (control) devices. Device design, performance projection and trade-off were presented, thus predicting the possible maximum performance in the strained-Si devices. The authors have also presented an analysis and design of strained-Si devices and circuits for low-power applications. Main focus has been on the static and dynamic power analysis for strained-Si CMOS. To study strained-Si CMOS logic circuits, nine-stage CMOS inverter ring oscillators were simulated via physics and process-based UFPDB model [44]. The model parameters were calibrated against the fabricated 70-nm devices where the same process condition is used for bulk-Si (control) and strained-Si devices.

Fig. 8.20 shows (fan out of 1) unloaded nine-stage ring oscillator delays vs. $V_{dd}$ for the bulk-SS and control CMOS circuits. The bulk strained-Si speed relative to the control one is significantly improved for equal $V_{dd}$-$V_t$ where $V_{dd}$ is 0.2 V higher for the control CMOS devices than the bulk strained-Si ones. Fig. 8.21 compares predicted (fan out of 1) unloaded nine-stage ring-oscillator delays for SSOI p-MOSFET with 20%-enhanced low-field hole mobility to that with 40%-enhanced low-field hole mobility. Fig. 8.22 shows SSOI CMOS ring-oscillator (four fundamental) delays as a function of the device width ratio ($\beta$) at $V_{dd}$ = 1.2 V. Fig. 8.23 shows predicted (fan out of 1) loaded nine-stage ring oscillator delays at $V_{dd}$ = 1.2 V. The SSOI CMOS exhibits significant speed improvement for $C_L$ = 0-100 fF over the bulk-SS counterpart, due in part to the reduced $C_J$ and in part to the forward-biased $V_{BS}$ and capacitive coupling effect in the floating-body configuration. It is noted that the SSOI delay is far superior to the bulk-Si. Predicted delay vs. power-delay product for (fan out of 1) unloaded inverter ring oscillator for bulk-Si, bulk-SS, and SSOI devices is shown in Fig. 8.24.

Fig. 8.25 compares predicted results of energy dissipation over one switching cycle for the bulk-Si, bulk-SS, and SSOI CMOS circuits. The three devices have equal $L_{eff}$ = 70 nm, equal $t_{ox}$ = 2.2 nm, and same $W_n/W_p$ = 10 $\mu$m/20 $\mu$m. The equal $C_{ov}$ is assumed for the three CMOS devices. The $V_t$'s for the three devices are the same at low $V_{ds}$. The SSOI CMOS exhibits the lowest energy dissipation primly due to much less $C_J$. The variation of the energy

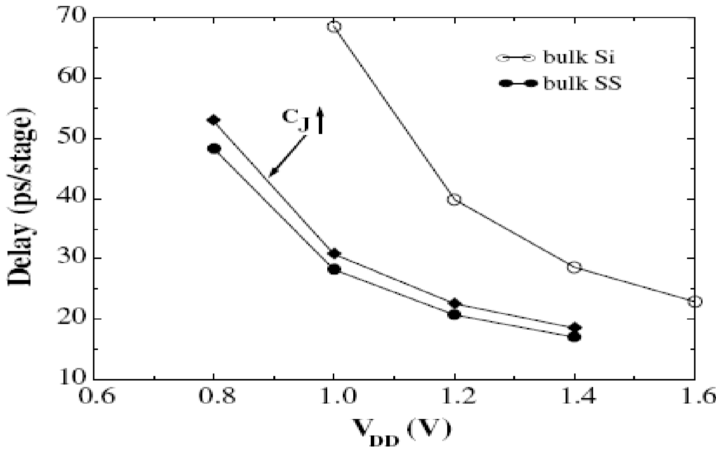



**FIGURE 8.20**
Predicted delays for (fan out of 1) unloaded nine-stage inverter ring oscillator. Case for the bulk-SS CMOS with higher $C_J$ is also shown. Since $V_t$ is 0.2 V higher for the bulk-Si devices, $V_{dd}$ is raised by 0.2 V. After K. Kim et al., *IEEE Intl. SOI Conf. Dig.*, 2002(17-19). With permission.

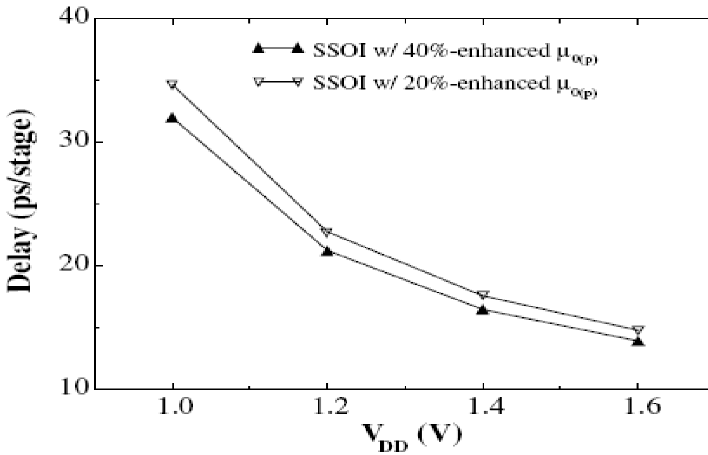



**FIGURE 8.21**
Predicted delays for (fan out of 1) unloaded nine-stage SSOI CMOS inverter ring oscillator. After K. Kim et al., *IEEE Intl. SOI Conf. Dig.*, 2002(17-19). With permission.

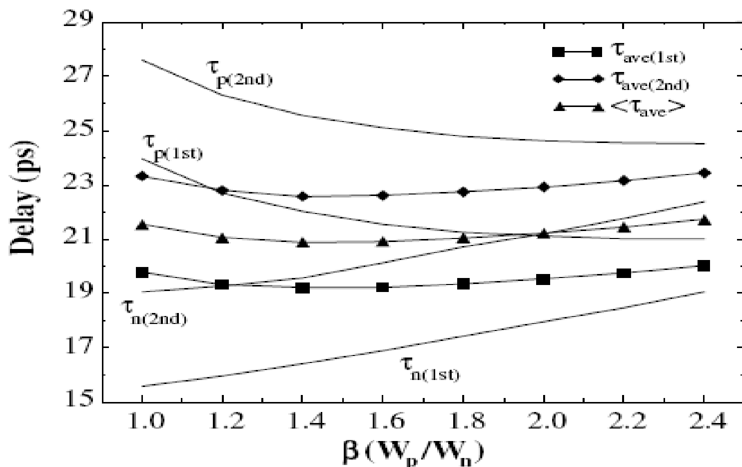



**FIGURE 8.22**
Predicted (four-fundamental) delays for (fan out of 1) unloaded nine-stage SSOI CMOS inverter ring oscillator at $V_{dd}$ = 2 V. After K. Kim et al., *IEEE Intl. SOI Conf. Dig.*, 2002(17-19). With permission.

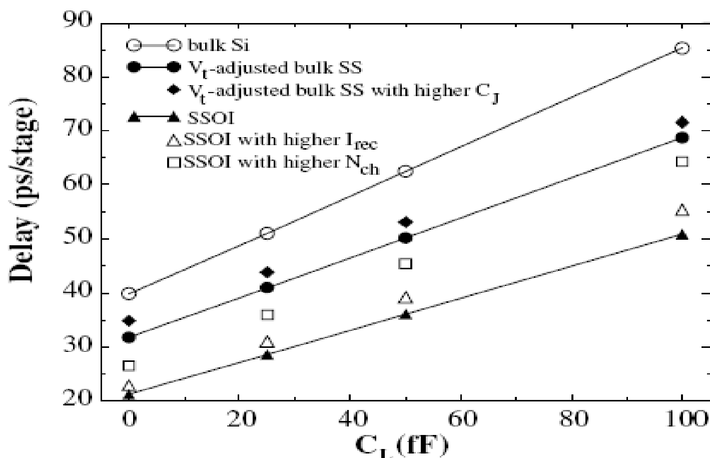



**FIGURE 8.23**
Predicted delays vs. load capacitance for (fanout of 1) nine-stage inverter ring oscillator. Cases for the $V_t$-adjusted bulk-SS and SSOI CMOS devices are shown. SSOI CMOS design point and performance trade-off are shown as well. After K. Kim et al., *IEEE Intl. SOI Conf. Dig.*, 2002(17-19). With permission.

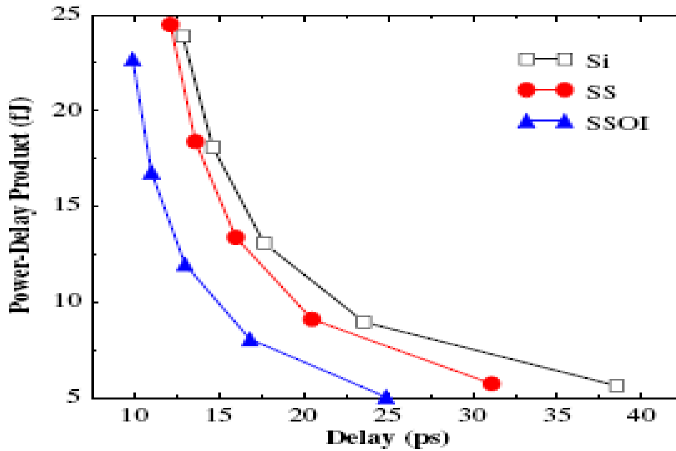



**FIGURE 8.24**
Predicted delay vs. power-delay product for (fan out of 1) unloaded inverter ring oscillator for bulk-Si, bulk-SS, and SSOI devices. After K. Kim et al., *IEEE Intl. SOI Conf. Dig.*, 2002(17-19). With permission.

dissipation is predicted within 5% for the SSOI CMOS with $V_t$ adjusted via increasing $I_{rec}$.

Thus, based on theoretical studies and hardware-calibrated models, strained-Si CMOS would offer lower static power consumption by ∼ 5-45% and dynamic power consumption by ∼ 10-28% for typical circuits. Furthermore, SSOI, building strained-Si channel with SOI structure, achieves additional performance improvement by ∼ 10% and lower power-delay product by ∼ 33% than bulk strained-Si counterpart for the inverters. However, due to the lower bandgap in strained-Si, SSOI could have different floating-body effect compared with SOI since all the generation and recombination currents are different, therefore, $V_t$ variation would be different for SSOI circuits. Furthermore, SSOI could have more severe self-heating effects than SOI due to the lower thermal conductivity in SiGe layer.

## 8.8 Summary

CMOS devices that feature strained-Si regions in the channel area are attracting significant interest as dramatic improvement of the electrical characteristics relative to conventional devices has been reported in strained channel CMOS transistors. The three main sources of stress in contemporary tech-

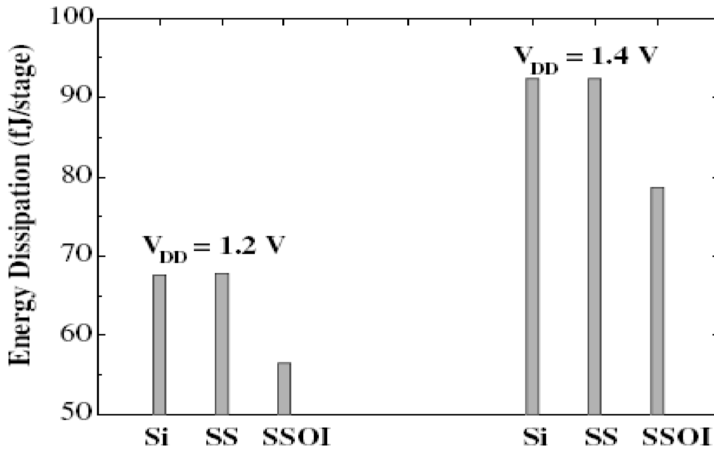



**FIGURE 8.25**
Predicted energy dissipations of (fan out of 1) unloaded nine-stage inverter ring-oscillator for bulk-Si, bulk-SS, and SSOI CMOS devices at $V_{dd} = 1.2$ V and 1.4 V. After K. Kim et al., *IEEE Intl. SOI Conf. Dig.*, 2002(17-19). With permission.

nologies are stresses due to thermal expansion, stresses due to material change, and stresses induced by external sources. It has become very important to understand the stress formation adequately in CMOS devices. TCAD simulation capabilities for strained-Si, both process and device, have been discussed in this chapter. From a user's perspective, the most important aspect of TCAD is its predictivity.

Ultimate circuit-level device model should include all the stress effects, including different channels and widths, S/D overhang regions, distance from adjacent active area in both longitudinal and transverse direction, and their sizes. Combination of different uniaxial stress from different directions will form biaxial stress effects, leading the model to be more complicated. The uniaxial channel stress is induced from the sides of the device. Devices with different channel lengths have different stress effects. Many mobility enhancement methods involve uniaxial stress through the active area, such as S/D area.

In this chapter, we have shown process-induced stress simulation in MOSFETs where a full 2D simulation including stresses was performed. The full stress history was simulated in order to account for the stresses induced. The change of mobility due to stress was taken into account in these simulations. The simulation shows how the stresses in the capping layer lead to tensile stress in the channel area. Several special material parameter models developed for the simulation of strained-Si and silicon-germanium structures were

presented. Extraction of BSIM3 SPICE parameters for strained-Si channel n-MOSFETs at room temperature has been presented.

## References

[1] D. Adams, M. Austin, R. Rai-Choudhury, and J. Hwang, "A performance comparison of advanced SOI technologies," *SOS/SOI Technology Workshop*, p. 79, 1988.

[2] J. H. Choi, Y. J. Park, and H. S. Min, "Electron mobility behavior in extremely thin SOI MOSFETs," *IEEE Electron Dev. Lett.*, vol. 16, pp. 527–529, 1995.

[3] S. Takagi, N. Sugiyama, T. Mizuno, T. Tezuka, and A. Kurobe, "Device structure and electrical characteristics of strained-Si-on-insulator (strained-SOI) MOSFETs," *Mat. Sci. Eng. B*, vol. 89, pp. 426–434, 2002.

[4] A. Lochtefeld and D. A. Antoniadis, "On experimental determination of carrier velocity in deeply scaled NMOS: how close to the thermal limit?," *IEEE Electron Dev. Lett.*, vol. 22, pp. 95–97, 2001.

[5] M. V. Fischetti and S. E. Laux, "Long-range Coulomb interactions in small Si devices. Part I: performance and reliability," *J. Appl. Phys.*, vol. 89, pp. 1205–1230, 2001.

[6] F. Gamiz, J. B. Roldan, H. Kosina, and T. Grasser, "Improving strained-Si on $Si_{1-x}Ge_x$ deep submicron MOSFETs performance by means of a stepped doping profile," *IEEE Trans. Electron Dev.*, vol. 48, pp. 1878–1884, 2001.

[7] M. W. Fischetti and S. E. Laux, "Band structure, deformation potentials, and carrier mobility in strained-Si, Ge, and SiGe alloys.," *J. Appl. Phys.*, vol. 80, pp. 2234–2252, 1996.

[8] J. B. Roldan, F. Gamiz, J. A. Lopez-Villanueva, and J. E. Carceller, "A Monte Carlo study on the electron-transport properties of high performance strained-Si on relaxed $Si_{1-x}Ge_x$ channel MOSFETs," *J. Appl. Phys.*, vol. 80, pp. 5121–5128, 1996.

[9] G. A. Armstrong and C. K. Maiti, "Strained-Si channel heterojunction p-MOSFETs," *Solid-State Electron.*, vol. 487-498, pp. 487–498, 1998.

[10] C. Lombardi, S Manzini, A. Saporito, and M. Vanzi, "A physically based mobility model for numerical simulation of nonplanner devices," *IEEE Trans. Computer-Aided Design*, vol. 7, pp. 1164–1170, 1988.

[11] J. H. Davies, *The Physics of Low-Dimensional Semiconductors: An Introduction.* Cambridge University Press, New York, 1998.

[12] C. K. Maiti, L. K. Bera, and S. Chattopadhyay, "Strained-Si heterostructure field effect transistors," *Semicond. Sci. Technol.*, vol. 13, pp. 1225–1246, 1998.

[13] C. K. Maiti, N. B. Chakrabarti, and S. K. Ray, *Strained Silicon Heterostructures: Materials and Devices.* The IEE, UK, 2001.

[14] J. Welser, J. L. Hoyt, S. Takagi, and J. F. Gibbons, "Strain dependence of the performance enhancement in strained-Si n-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 373–376, 1994.

[15] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Fabrication and analysis of deep submicron strained-Si N-MOSFETs," *IEEE Trans. Electron Dev.*, vol. 47, pp. 1406–1415, 2000.

[16] D. K. Nayak, K. Goto, A. Yutani, J. Murota, and Y. Shiraki, "High-mobility strained-Si PMOSFETs," *IEEE Trans. Electron Dev.*, vol. 43, pp. 1709–1715, 1996.

[17] C. K. Maiti, L. K. Bera, S. S. Dey, D. K. Nayak, and N. B. Chakrabarti, "Hole mobility enhancement in strained-Si p-MOSFETs under high vertical fields," *Solid-State Electron.*, vol. 41, pp. 1863–1869, 1997.

[18] K. Rim, K. Chan, L. Shi, D. Boyd, J. Ott, N. Klymko, F. Cardone, L. Tai, S. Koester, M. Cobb, D. Canaperi, B. To, E. Duch, I. Babich, R. Carruthers, P. Saunders, G. Walker, Y. Zhang, M. Steen, and M. Ieon, "Fabrication and mobility characteristics of ultrathin strained-Si directly on insulator (SSDOI) MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 49–52, 2003.

[19] H. Yin, K. D. Hobart, R. L. Peterson, F. J. Kub, S. R. Shieh, T. S. Duffy, and J. C. Sturm, "Fully-depleted strained-Si on insulator NMOSFETs without relaxed SiGe buffers," *IEEE IEDM Tech. Dig.*, pp. 53–56, 2003.

[20] F. M. Bufler, Y. Asahi, H. Yoshimura, C. Zechner, A. Schenk, and W. Fichtner, "Monte Carlo simulation and measurement of nanoscale n-MOSFETs," *IEEE Trans. Electron Dev.*, vol. 50, pp. 418–424, 2003.

[21] F. M. Bufler and W. Fichtner, "Scaling of strained-Si n-MOSFETs into the ballistic regime and associated anisotropic effects," *IEEE Trans. Electron Dev.*, vol. 50, pp. 278–284, 2003.

[22] F. M. Bufler and W. Fichtner, "Scaling and strain dependence of nanoscale strained-Si p-MOSFET performance," *IEEE Trans. Electron Dev.*, vol. 50, pp. 2461–2466, 2003.

[23] S. K. Samanta, *Ultrathin Oxide and Oxynitride Ggate Dielectric Films for Silicon Hetero-FETs.* Ph.D. thesis, Indian Institute of Technology, Kharagpur, 2003.

[24] S. V. Vandebroek, E. Crabbe, B. Meyerson, D. Harame, P. Restle, J. Stork, and J. Johnson, "SiGe-channel heterojunction p-MOSFETs," *IEEE Trans. Electron Dev.*, vol. 41, pp. 90–101, 1994.

[25] A. J. Scholten and M. B. M. Klaassen, "Anomalous geometry dependence of source/drain resistance in narrow-width MOSFETs," *Proc. IEEE Int. Conf. on Microelectronic Test Structure*, pp. 77–82, 1998.

[26] P. I. Suciu and R. L. Johnston, "Experimental derivation of the source and drain resistance of MOS transistor," *IEEE Trans. Electron Dev.*, vol. 27, pp. 1846–1848, 1980.

[27] S. C. Sun and J. D. Plummer, "Electron mobility in inversion and accumulation layers on thermally oxidized silicon surfaces," *IEEE Trans. Electron Dev.*, vol. 27, pp. 1497–1508, 1980.

[28] S. Takagi, A. Toriumi, M. Iwase, and H. Tango,, "On the universality of inversion layer mobility in Si MOSFETs: Part I - Effects of substrate impurity concentration," *IEEE Trans. Electron Dev.*, vol. 41, pp. 2357–2362, 1994.

[29] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Transconductance enhancement in deep submicron strained-Si n-MOSFETs," *IEEE IEDM Tech. Dig.*, pp. 707–710, 1998.

[30] K. A. Jenkins, J. Y.-C. Sun, and J. Gautier, "Characteristics of SOI FETs under pulsed conditions," *IEEE Trans. Electron Dev.*, vol. 44, pp. 1923–1930, 1997.

[31] L. T. Su, J. E. Chung, A. D. Antoniadis, K. E. Goodson, and M. I. Flik, "Measurement and modeling of self-heating in SOI nMOSFETs," *IEEE Trans. Electron Dev.*, vol. 41, pp. 69–75, 1994.

[32] J. P. Dismukes, L. Ekstrom, E. F. Steigmeiter, I. Kudman, and D. S. Beers, "Thermal and elastic properties of heavily doped Ge-Si alloys up to 1300 K," *J. Appl. Phys.*, vol. 35, pp. 2899–2907, 1964.

[33] C. K. Maiti and G. A. Armstrong, *Applications of Silicon-Germanium Heterostructure Devices.* Institute of Physics Publishing, UK, 2001.

[34] K. A. Jenkins and K. Rim, "Measurement of the effect of self-heating in strained-silicon MOSFETs," *IEEE Electron Dev. Lett.*, vol. 23, pp. 360–362, 2002.

[35] S. H. Olsen, A. G. O'Neill, L. S. Driscoll, K. S. K. Kwa, S. Chattopadhyay, A. M. Waite, Y. T. Tang, A. G. R. Evans, D. J. Norris, A. G. Cullis, D. J. Paul, and D. J. Robbins, "High-performance nMOSFETs using a novel strained Si/SiGe CMOS architecture," *IEEE Trans. Electron Dev.*, vol. 50, pp. 1961–1969, 2003.

[36] N. D. Jankovic, T. V. Pesic, and A. O'Neill, "Modelling of strained-Si/SiGe NMOS transistors including DC self-heating," *Solid-State Electron.*, vol. 50, pp. 496–499, 2006.

[37] Synopsys, "SENTAURUS TCAD Tools," 2005.

[38] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors." San Jose, Calif., 2005.

[39] ISE-TCAD, Switzerland, "SPARTA Manual," 2004.

[40] Synopsys, "TCAD News," October 2005.

[41] K. Kim, C.-T. Chuang, K. Rim, and R. V. Joshi, "Performance assessment of scaled strained-Si channel-on-insulator (SSOI) CMOS," *IEEE Intl. SOI Conf. Dig.*, pp. 17–19, 2002.

[42] K. Kim, C.-T. Chuang, K. Rim, and R. V. Joshi, "Performance assessment of scaled strained-Si channel-on-insulator (SSOI) CMOS," *Solid-State Electron.*, vol. 48, pp. 239–243, 2004.

[43] K. Kim, R. V. Joshi, and C.-T. Chuang, "Power analysis of strained-Si devices/circuits," *Solid-State Electron.*, vol. 48, pp. 1453–1460, 2004.

[44] J. G. Fossum and W. Zhang, "Performance projections of scaled CMOS devices and circuits with strained Si-on-SiGe channels," *IEEE Trans. Electron Dev.*, vol. 50, pp. 1042–1049, 2003.

# Index


activation energy, 69
alloy scattering, 73, 132
AMD, 367
ANSYS, 409
ATLAS, 247, 261, 268, 387

bandgap narrowing, 130
biaxial strain, 315
Boltzmann transport equation, 386
Boltzmann-Matano analysis, 197

CMOS, 386, 400
conduction band discontinuity, 248
Coulomb scattering, 324

$\delta$-doping, 251
dielectric
- breakdown, 3, 202
- constant, 221

engineered-substrates, 12, 59, 71, 106
- electronic properties, 119
- technology, 64

engineering
- channel, 306
- drain, 306
- gate, 281
- mobility, 406
- source, 306
- substrate, 8

epitaxy, 60

FDSOI, 14

gate dielectrics, 12
- on engineered-substrates, 189

gate oxide
- characterization
  - chemical analysis, 226
  - conduction mechanism, 230
- effect of
  - layer thickness, 210
  - surface roughness, 207
- reliability, 232

generation and recombination, 211, 415
GeOI, 90
GeOI substrate, 90

HEMTs
- SiGe-based, 277
  - design issues, 281
  - layer design, 286

hetero-FETs
- characterization, 399
- modeling, 385
- simulation, 385
- technology CAD, 400

heteroepitaxy, 64
high-k gate dielectrics
- on Ge, 234
- on strained-Si, 220

hot-carrier, 200, 250, 333
hybrid orientation technology (HOT), 42
hybrid substrate, 89
hydrodynamic model, 249

impurity scattering, 131, 279, 385, 404
interface state density, 190, 206, 211, 251, 342
interface traps, 189, 197, 211

lattice constant, 59, 65

lattice scattering, 131

Matthews and Blakeslee theory, 69
mechanical stress control, 43
metal-oxide-semiconductor (MOS), 189
microwave plasma deposition, 225
mobility, 11, 23, 106
  mobility engineering, 11, 406
  orientation dependence, 46, 341
  temperature dependence, 157
  thickness dependence, 150
MODFET, 7, 277
Moore's law, 1
MOS capacitor, 149, 201, 210
MOSFET
  buried channel, 73, 199, 245, 250
  design issues, 247
  double-gate, 343
  SiGe/SiGeC, 245, 273
    operation, 250
  single-gate, 343
  surface channel, 250, 257
  ultrathin body (UTB), 15, 386

nitride-capping layer, 119
nonclassical CMOS structures, 14

phonon scattering, 110, 132
plasma deposition, 225
plasma nitridation, 203
plasma oxidation, 189
poly-SiGe, 250
process integration, 13
process-induced stress, 10, 20, 37, 176

scattering mechanism, 138
setback layer, 392
Si-isotope, 109
$Si_{1-x}Ge_x$
  oxidation, 198
SiGe MOSFET
  on SOI, 267
SiGe/SiGeC
  material parameters, 247
silicide, 28
Smart-Cut, 97
SPICE parameter, 409
  extraction, 409
SSDOI, 10, 390
SSOI, 10, 43, 59, 77
strain
  global, 7, 25, 32, 322
  local, 7, 23, 25, 32, 309
  process-induced, 7, 10, 23, 33, 37, 120, 160, 176, 295
  substrate-induced, 9, 11, 35, 46, 120, 176
strain analysis, 48
strain compensation, 72, 273
strain control, 64
strain engineering, 8, 9, 23
  biaxial, 176
  uniaxial, 176
strain relaxation, 303, 322, 366
strain-engineered MOSFET
  modeling, 16, 385
  simulation, 16, 385
strain-engineered substrates, 59
strained-Ge, 48, 77, 199
strained-layer
  deposition
    GSBME, 76
    MBE, 74
    RPCVD, 191
    RTCVD, 74
    SiGe, 245
    SiGeC, 245
    UHVCVD, 76
strained-layer characterization
  AFM, 95
  Raman, 101
  RBS, 94
  reciprocal lattice mapping, 94
  SIMS, 100
  TEM, 28, 323
  XRD, 97
strained-Si

(process-induced), 160
electron mobility, 170
hole mobility, 166
(substrate-induced), 120
band structure, 122
electron mobility, 134
energy gap, 122
hole mobility, 138, 145, 146
band alignment
type-I, 73
type-II, 73
band offset, 73
band structure, 9
bandgap, 71
compositional grading, 74
critical layer thickness, 77
dislocation density, 74
layer formation, 74
misfit dislocation, 74
MOSFET
performance, 412
nitridation, 203
on-insulator, 81, 203
rapid thermal oxidation, 200
relaxed-SiGe layer, 74
scaling, 323
SiC-stressor, 309
SiGe-stressor, 306
specification, 80
subband structure, 47
substrate, 74
technology, 13
technology CAD, 400
thermal conductivity, 109
thermal oxidation, 191
Ge outdiffusion, 193
kinetics, 197
strained-Si MOSFET
dual channel, 346
layout dependence, 333
operating principle, 296
orientation dependence, 341
reliability, 353
self-heating, 358
structures, 190
surface channel, 190
thickness dependence, 338
threshold voltage, 297
stress
analysis, 48
contour, 41, 329
during manufacturing, 16
substrate engineering, 5, 10, 92
substrate-induced strain, 10, 20, 37, 176
superlattice, 62, 123
surface scattering, 250, 270

tensile stress, 119
thermal stability, 30, 72
TSMC, 360

uniaxial stress, 306

valence band discontinuity, 270, 398
valence band offset, 130
Vegard's law, 28, 65
velocity overshoot, 141, 286
velocity saturation, 120, 138
virtual substrate, 74, 338