Numerical Simulation of Sub-100 nm Strained Si/SiGe MOSFETs for RF and CMOS Applications

Lianfeng Yang

A thesis submitted to the Department of Electronics and Electrical Engineering, Faculty of Engineering at the University of Glasgow

for the degree of

Doctor of Philosophy

UNIVERSITY
of
GLASGOW

Device Modelling Group

Department of Electronics and Electrical Engineering

University of Glasgow

United Kingdom

September 2004

Declaration

The work in this thesis is based on research carried out at the Device Modelling Group, University of Glasgow, UK. No part of this thesis has been submitted elsewhere for any other degree or qualification and it all my own work unless referenced to the contrary in the text.

Dedicated to

my parents and my wife .....

献给

父亲杨承华、母亲陈德存

和妻子高毓航

Abstract

Drift-Diffusion, Hydrodynamic and Monte Carlo simulations have been used in this work to simulate strained Si/SiGe devices for RF and CMOS applications. For numerical simulations of Si/SiGe devices, strain effects on the band structure of Si have been analyzed and analytical expressions are presented for parameters related to the bandgap and band alignment of Si/SiGe heterostructure.

Optimization of n-type buried strained Si channel Si/SiGe MODFETs has been carried out in order to achieve high RF performance and high linearity. The impact of both lateral and vertical device geometries and different doping strategies has been investigated.

The impact of the Ge content of the SiGe buffer on the performance of p-type surface channel strained Si/SiGe MOSFETs has been studied. Hydrodynamic device simulations have been used to assess the device performance of p-type strained Si/SiGe MOSFETs down to 35 nm gate lengths. Well-tempered strained Si MOSFETs with halo implants around the source/drain regions have been simulated and compared with those devices possessing only a single retrograde channel doping.

The calibrations in respect of sub-100 nm Si and strained Si MOSFETs fabricated by IBM lead to a scaling study of those devices at 65 nm, 45 nm and 35 nm gate lengths. Using Drift-Diffusion simulations, ring oscillator circuit behaviour has been evaluated. Strained Si on insulator (SSOI) circuits have also been simulated and compared with strained Si circuits, Si circuits employing conventional surface channel MOSFETs along with SOI devices.

Ensemble Monte Carlo simulations have been used to evaluate the device performance of n-type strained Si MOSFETs. A non-perturbative interface roughness scattering model has been used and validated by calibrating with respect to ex-

perimental mobility behaviour and device characteristics. The impact of interface roughness on the performance enhancement of strained Si MOSFETs has been investigated and evidence for reduced interface roughness scattering is presented, *i.e.*, a smoother interface is suggested in strained Si MOSFETs. A 35 nm gate length Toshiba Si MOSFET has been simulated and the performance enhancement of 35 nm strained Si MOSFETs over the Toshiba Si device is predicted.

Monte Carlo simulations are also employed to investigate the performance degradation due to soft-optical phonon scattering, which arises with the introduction of high-κ gate dielectrics. Based on the device structures of the calibrated sub-100 nm n-type conventional and strained Si IBM MOSFETs, significant current degradation has been observed in devices with high-κ gate dielectrics, HfO_2 and Al_2O_3.

Acknowledgements

I express my most sincere gratitude to my supervisor, Professor John R. Barker for the encouragement, guidance and support he has given me throughout my stay at Glasgow. I also express my most sincere gratitude to Professor Asen Asenov for his technical guidance and encouragement.

I had the great pleasure of working very closely with Dr. Jeremy R. Watling and I express my most sincere gratitude for his guidance and insights into MOSFET Monte Carlo simulations.

I would like to acknowledge Mr. Richard C. W. Wilkins for all his help throughout my stay at Glasgow, especially his help in correcting most of my English writing and sharing his ideas of surface roughness scattering.

I would like to thank Dr. Scott Roy and Dr. Mirela Boriçi, who have been working in the same project with me, for their help and sharing many of their ideas with me. I also would like to acknowledge Dr. Karol Kalna for the frequent discussions in various areas and the collaborations on the III-V MOSFET simulations. I also express my gratitude to Dr. Binjie Cheng, who has worked in the same office and shared lots of information and fun with me for two years .

I also owe a great deal to all members of the Device Modelling Group for the fun time we had working together. I would especially like to thank Mr. Andrew Brown, Dr. Campbell Miller, Mr. Gareth Roy, Mr. Craig Alexander, Ms. Angelica Lee, Mr. Yoshiaki Kitano, and Dr. Antonio Martinez. I especially thank Mr. Fikru Adam-Lema for sharing information on 35 nm MOSFET simulations.

I would like to thank Dr. Khaled Elgaid and Professor Iain Thayne from the Ultrafast System Group for sharing information on the 70 nm Si/SiGe MODFET and the device design for high linearity applications. I also wish to thank Dr. Thomas

Hackbarth from the Daimler Chrysler for supplying experimental data and sharing information of the Si/SiGe MODFETs for RF applications.

I would like to thank the guys from computer support, Mr. William (Billy) Allan and Mr. John Trinder, for their continual and efficient technical support.

I would like to acknowledge all my friends here in Glasgow. I have enjoyed the last three years of my life here with them and I wish to keep in touch with all of them in the future.

I would like to thank the Engineering and Physical Sciences Research Council (EPSRC) for their full financial support of this work under the grant number: GR/N65677/01, "SiGe for MOS technologies: Phase II". I also would like to acknowledge IBM for providing computer resources under a Shared University Research Grant (SUR) and Synopsys for supporting their TCAD software suite. I would also like to thank the Department of Electronics and Electrical Engineering and the University of Glasgow for the working environment and facilities.

Finally, I am deeply indebted to my family - my parents, parents-in-law, brother and my wife, Yuhang (Hazel) Gao for their love and support throughout my seemingly never-ending studies. Indeed, without the understanding and continual encouragement from my wife, I could not have finished this work. This work is dedicated to them.

Publications and Conferences

Journal papers

1. **L. Yang**, J. R. Watling, R. C. W. Wilkins, M. Boriçi, A. Asenov and J. R. Barker. Surface roughness scattering in sub-100nm Si and strained Si MOSFETs - A Monte Carlo study. submitted to *IEEE Transaction on Electron Devices*, 2004.

2. **L. Yang**, J. R. Watling, R. C. W. Wilkins, M. Boriçi, J. R. Barker, A. Asenov and S. Roy. Si/SiGe Heterostructure Parameters for Device Simulations. *Semiconductor Science and Technology*, 19(10):1174-1182, 2004.

3. **L. Yang**, A. Asenov, J. R. Watling, M. Boriçi, J. R. Barker, S. Roy, K .Elgaid, I. Thayne and T. Hackbarth. Impact of device geometry and doping strategy on linearity and RF performance in Si/SiGe MODFETs. *Microelectronics Reliability*, 44(7):1101-1107, 2004.

4. J. R. Watling, **L. Yang**, M. Boriçi, R. C. W. Wilkins, A. Asenov, J. R. Barker and S. Roy. The impact of interface roughness scattering and degeneracy in relaxed and strained Si n-channel MOSFETs. *Solid State Electronics*, 48:1337-1346, 2004 (invited).

5. K. Kalna, M. Boriçi, **L. Yang** and A. Asenov. Monte Carlo simulation of III-V MOSFETs. *Semiconductor Science and Technology*, 19(4):S202-205, 2004.

6. M. Boriçi, J. R. Watling, R. C. W. Wilkins, **L. Yang**, J. R. Barker and A. Asenov. Interface roughness scattering and its impact on electrons transport in

relaxed and strained Si n-MOSFETs. *Semiconductor Science and Technology*, 19(4):S155-157, 2004.

7. **L. Yang**, J. Watling, M. Boriçi, R. C. W. Wilkins, A. Asenov, J. R. Barker and S. Roy. Simulations of scaled sub-100nm strained Si/SiGe p-channel MOSFETs. *Journal of Computational Electronics*, 2(4):363-368, 2003.

8. J. Watling, **L. Yang**, M. Boriçi, J. R. Barker and A. Asenov. Degeneracy and high doping effects in deep sub-micron relaxed and strained Si n-MOSFETs. *Journal of Computational Electronics*, 2(4):475-480, 2003.

9. K. Kalna, **L. Yang** and A. Asenov. Simulation study of high performance III-V MOSFETs for Digital Applications. *Journal of Computational Electronics*, 2(4):341-346, 2003.

10. M. Boriçi, J. R. Watling, R. C. W. Wilkins, **L. Yang** and J. R. Barker. A Non perturbative model of surface roughness scattering for Monte Carlo simulation of relaxed Silicon n-MOSFETs. *Journal of Computational Electronics*, 2(4):163-168, 2004.

Conference talks and posters

1. **L. Yang**, J. R. Watling, F. Adam-Lema, A. Asenov and J. R. Barker. Scaling study of Si and strained Si n-MOSFETs with different gate stacks. to appear in the *2004 IEEE International Electron Devices Meeting (IEDM)*, San Francisco, 2004.

2. J. R. Watling, **L. Yang**, A. Asenov, J. R. Barker and S. Roy. Impact of high-κ dielectric HfO_2 on the mobility and device performance of sub-100 nm n-MOSFETs. to appear in the *2004 International Workshop on Electrical Characterization and Reliability of High-κ* Devices , Austin, 2004. (to be published in the *IEEE Transaction on Device and Material Reliability*)

3. **L. Yang**, J. R. Watling, F. Adam-Lema, A. Asenov and J. R. Barker. Simulations of scaled strained Si MOSFETs with high-κ gate stacks. to appear in

the *10th IEEE International Workshop of Computational Electronics (IWCE)*, Indiana, 2004. (to be published in the *Journal of Computational Electronics*)

4. **L. Yang**, J. R. Watling, A. Asenov, J. R. Barker and S. Roy. Sub-100nm strained Si CMOS: Device performance and circuit behavior. to appear in the *7th International Conference on Solid-State and Integrated-Circuit Technology (ICSICT)*, Beijing, 2004.

5. **L. Yang**, J. R. Watling, A. Asenov, J. R. Barker and S. Roy. Device performance in conventional and strained Si MOSFETs with high-κ gate stack. *Proceedings of the 2004 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 199-202, Springer-Velag, 2004.

6. **L. Yang**, J. R. Watling, J. R. Barker and A. Asenov. Impact of soft-optical phonon scattering due to high-κ dielectrics on the performance of sub-100nm conventional and strained Si n-MOSFETs. presented at the *27th International Conference on Physics of Semiconductors (ICPS-27)*, Arizona, 2004 (to be published by the American Institute of Physics).

7. **L. Yang**, J. R. Watling, R. C. W. Wilkins, J. R. Barker and A. Asenov. Monte Carlo investigation of 'reduced' interface roughness in sub-100nm strained Si MOSFETs. *Proceedings of the 5th Europe Workshop on Ultimate Integration of Silicon (ULIS04)*, pages 23-26, IMEC, 2004.

8. K. Kalna, **L. Yang**, J. R. Watling and A. Asenov. 80nm InGaAs MOSFET compared to equivalent Si transistor. *Proceedings of the 5th Europe Workshop on Ultimate Integration of Silicon (ULIS04)*, pages 159-162, IMEC, 2004.

9. **L. Yang**, J. R. Watling, R. C. W. Wilkins, J. R. Barker and A. Asenov. Monte Carlo investigation of interface roughness scattering in relaxed and strained Si n-MOSFETs. presented at the *IoP Condensed Matter and Materials Physics Conference (CMMP04)*, Warwick, 2004.

10. J.Watling, **L. Yang**, J. R. Barker and A. Asenov. The Impact of high-κ dielectrics on the future performance of nano-scale MOSFETs. presented at the

IoP Condensed Matter and Materials Physics Conference (CMMP04), Warwick, 2004.

11. **L. Yang**, A. Asenov, M. Boriçi, J. R. Watling, J. R. Barker, S. Roy, K .Elgaid, I. Thayne and T. Hackbarth. Optimizations of sub-100nm Si/SiGe MODFETs for high linearity RF applications. *Proceedings of the 2003 IEEE Conference on Electron Device and Solid-State Circuits (EDSSC03)*, pages 331-334, Hong Kong, 2003.

12. **L. Yang**, A. Asenov, J. R. Watling, M. Boriçi, J. R. Barker, S. Roy, K .Elgaid, I. Thayne and T. Hackbarth. A study of high linearity Si/SiGe HFETs. *Proceedings of the 14th Workshop on Modelling and Simulation of Electron Device (MSED03)*, pages 41-44, Barcelona, 2003.

13. **L. Yang**, J. Watling, M. Boriçi, R. C. W. Wilkins, A. Asenov, J. R. Barker and S. Roy. Simulations of scaled sub-100nm strained Si/SiGe p-channel MOSFETs. presented at the *9th IEEE International Workshop of Computational Electronics (IWCE)*, Rome, 2003.

14. **L. Yang**, J. R. Watling, R. C. W. Wilkins, A. Asenov, J. R. Barker, S. Roy and T. Hackbarth. Scaling study of Si/SiGe MODFETs for RF applications. presented at the *10th Advanced Heterostructure Workshop (AHW02)*, Hawaii, 2002.

15. **L. Yang**, J. R. Watling, R. C. W. Wilkins, A. Asenov, J. R. Barker, S. Roy and T. Hackbarth. Scaling study of Si/SiGe MODFETs for RF applications. *Proceeding of the 10th IEEE International Symposium on Electron Devices for Microwave and Optoelectronic Applications (EDMO02)*, pages 101-106, Manchester, 2002.

16. K. Kalna, **L. Yang** and A. Asenov. High Performance III-V MOSFETs: a Dream Close to Reality?. *Proceeding of the 10th IEEE International Symposium on Electron Devices for Microwave and Optoelectronic Applications (EDMO02)*, pages 243-248, Manchester, 2002.

Glossary

ϵ_0	Vacuum permittivity, 8.85×10^{-14} F/cm
f_T	Cut-off frequency
f_{max}	Maximum frequency of oscillation
g_{ds}	Output conductance (dI_D/dV_D)
g_m	Transconductance (dI_D/dV_G)
$\hbar$	Reduced Planck's constant, 1.055×10^{-34} J-s
k_B	Boltzmann's constant, 1.38×10^{-23} J/K
m_0	Free electron mass, 9.1×10^{-31} kg
q	Electronic charge, 1.6×10^{-19} C
E_c	Conduction band
E_g	Band gap
E_v	Valence band
I_D	Drain current
L_{eff}	Effective gate length
L_g	Physical gate length
N_A	Acceptor (doping) concentration
N_D	Donor (doping) concentration
S	Subthreshold slope
V_D	Drain voltage
V_G	Gate voltage
V_T	Threshold voltage
1-D (1D)	One Dimensional
2-D (2D)	Two Dimensional
2-DEG	2-Dimensional Electron Gas

BTE	Boltzmann Transport Equation
CL (λ)	Correlation Length (of surface roughness)
CMOS	Complemented Metal Oxide Semiconductor
DCFET	Doped Channel Field Effect Transistor
DDM	Drift-Diffusion (transport) Model
DIBL	Drain Induced Barrier Lowering
DOS	Density of States
EOT	Equivalent Oxide Thickness
FD	Fully Depleted (SOI)
HDM	Hydrodynamic (transport) Model
HEMT	High Electron Mobility Transistor
HFET	Heterostructure Field Effect Transistor
IM	Intermodulation
IR	Interface Roughness
ITRS	International Technology Roadmap for Semiconductors
LDD	Lightly doped drain
LO	Longitudinal Optical (phonon)
MC	Monte Carlo
ME	Modulation Efficiency
MODFET	Modulation Doped Field Effect Transistor
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
PD	Partially Depleted (SOI)
PIP3	Input power leading to excess third-order intermodulation
RF	Radio Frequency
RMS (Δ)	Root Mean Square (height of surface roughness)
SCE	Short Channel Effects
S/D	Source/drain
SDE	Source Drain Extension
SGOI	SiGe-on-insulator
SIMOX	Separation by implanted oxygen
SO	Soft Optical (phonon)

SOI	Si-on-Insulator MOSFET
SSDOI	Strained Si directly on Insulator MOSFET
SSi	Strained Si MOSFET
SSOI	Strained Si on SiGe SOI MOSFET
TO	Transverse Optical (phonon)

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Silicon based CMOS technology is now entering the sub-100 nm regime [1], enabling the integration of tens of millions transistors onto a single chip, while the dimensions of the smallest demonstrated devices [2, 3] are rapidly approaching their 'physical' limit - the silicon lattice constant. However, this continuous dimension scaling not only brings difficulties to device fabrication, but also induces device operation problems. The major problems of scaling conventional MOS-type devices include [4]: (1) Quantum mechanical tunnelling through the thin gate oxide, from source to drain and from drain to body; (2) threshold voltage control induced by random doping effects; (3) short channel effects and mobility degradation; (4) process control of thin layer uniformity, accurate lithography and implantation.

Therefore, to extend the lifetime of the silicon based CMOS technology, devices with new structures or new materials need to be considered [1, 4]. In a scaled conventional Si MOSFET, a thin gate oxide and a highly doped channel are required in order to have a good control of short channel effects (SCE). However, such action induces problems of gate tunnelling and low mobility in the channel. Complicated designs of source/drain (S/D) and well doping profiles, *e.g.*, lightly doped drain (LDD), retrograde or halo doping, are necessary to suppress the serious SCE and sustain channel mobility, maintaining a high on/off current ratio. The use of high-κ gate dielectrics [5] is a solution to suppress the gate leakage, although problems still exist in terms of process control and associated mobility degradation [6, 7]. The evolution of device structure [1, 4], *e.g.*, ultra-thin body SOI MOSFET and

multi-gate MOSFETs with lightly doped or undoped channel, which benefit from intrinsically well controlled SCE and better carrier transport, are promising solutions for future ultra-scaled devices. The scaling also necessiates source/drain engineering by using either Schottky source/drain or non-overlapped source/drain extensions in order to reduce source/drain resistance or suppress 2-D effects. These performance boosters have been summarized in the 2003 edition of the International Technology Roadmap for Semiconductors (ITRS) [1] and the roadmap suggested that one or several boosters may be required for devices beyond the 90 nm technology node in order to sustain the historic annual increase of intrinsic speed of high-performance MPUs at 17% [1]. Among those potential solutions for future CMOS applications, transport enhanced FETs using strained Si channel are amongst the most mature technology and have already demonstrated enhanced device and circuit performance.

1.1 The Allure of Strained Si

The transport enhanced FETs use new materials with higher mobilities to enhance/replace the Si channel. Various materials have been proposed, *e.g.*, III-V materials, Ge or SiGe, and strained Si. III-V materials have higher mobilities than bulk Si, but they lack a natural oxide like SiO_2 in the case of Si. Recent research has shown that growing dielectrics on III-V materials with good interfaces is achievable, indicating the possible future of III-V material for CMOS applications [8]. On the other hand, although Ge or SiGe channel with higher mobilities have been studied primarily for p-type applications, the difficulties of growing an insulator with good interface properties on top of Ge or SiGe has slowed down its application for CMOS [9, 10]. Recent research focuses are on the strained Si material, which has advantages in terms of the mobility enhancement for both electrons and holes and its compatibility with existing Si technology.

Strained Si may be achieved either through process induced material stress or by pseudomorphically growing silicon on a buffer with larger lattice constant, usually relaxed SiGe. Appropriate process steps can stress the Si channel through appropriate strain engineering of shallow trench isolation (STI), cap layer (contact-etch-stop

layer) or silicidation processes [11–13], which has been incorporated into Intel's current 90 nm technology to achieve performance enhancement [11]. However, this kind of strain is very dependent on the specific process conditions and also device dimensions [12]. In comparison with the process induced strain, strained Si pseudomorphically grown on a relaxed SiGe layer is a commonly adopted approach to achieve tensile strain. If the strain Si layer is thinner than its *critical thickness* [14] (defined as the thickness beyond which the strained Si can be grown without inducing misfit dislocations to relieve the strain), it should be reasonably stable.

Strained Si on SiGe heterostructure for CMOS applications has been studied for more than ten years [15] and IBM first announced its plan in 2001 to use strained Si for their future CMOS technology [16], which was claimed to boost the chip speed by 35% with only a 10% cost increase. Compared to bulk Si CMOS, a 70 nm strained Si process has been recently demonstrated, delivering a 95% higher inverter peak current and a 2.2 ps reduction in ring oscillator delay for the same drive current [17]. 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 [4,18] and the strained Si directly on insulator (SSDOI) MOSFET [19,20]. Having a highly strained Si channel [21] or using a different orientation (110) substrate [22] in p-type MOSFETs, the performance match between the NMOS and PMOS for CMOS applications might be achieved. However, unless specifically indicated, all the work contained in this thesis refer to the (100) substrate.

Driven by the requirements for high performance and high integration, strained Si has been widely accepted as a candidate for future CMOS applications [1] and sub-100 nm strained Si devices have been demonstrated by major industry vendors such as IBM [16], Intel [11], AMD [23], Toshiba [18], UMC [17], TSMC [24].

Apart from its wide application in CMOS, strained Si/SiGe heterostructure has also been applied to the high electron mobility transistor (HEMT) structure, in which the undoped strained Si channel is capped by two relaxed SiGe layers and the carriers are supplied by remote doped SiGe layers, forming a high mobility 2-D quantum well. This structure, named the modulation-doped field effect transistor (MODFET), has demonstrated promising RF performance of highest record f_{max},

188 GHz [25], and minimum noise figures of 0.3 dB at 25 GHz [26]. Compared to their III-V counterparts, strained Si MODFET is more easily integrated into existing Si technology and is therefore able to realize the system-on-chip in some application areas.

1.2 About the Project

The Ph.D project is part of "SiGe for MOS Technologies - Phase II", a large inter-university collaborative grant funded by the UK Engineering Physical and Science Research Council (EPSRC). The work has been carried out within the Device Modelling Group at the Department of Electronics and Electrical Engineering of the University of Glasgow. The aim of the project is to study the evolution of Si based MOSFETs via the incorporation of new materials, for example SiGe or strained Si, to evaluate the resultant device performance and predict future scaling trends of strained Si/SiGe MOSFETs for CMOS and RF applications. The project is based on simulation work, using the commercial TCAD tools MEDICI and TAURUS from Synopsys [27]; a 1-D Poisson-Schrödinger solver [28] and an in-house ensemble Monte Carlo simulator [29].

The first year's work focused on the scaling study of n-type strained Si/SiGe MODFETs [30], in collaboration with DaimlerChrysler in Germany. A simulation based calibration with respect to the experimental data of 67 nm effective gate length n-type bulk Si and strained Si MOSFETs from IBM [31] was also carried out during this period.

In the second year, further collaboration with Daimler Chrysler and the Ultrafast System Group in the department led to the continuation of the work on strained Si/SiGe MODFETs, focusing on the optimization of the device structure for high linearity RF applications [32, 33]. Major work in this year was the scaling study of sub-100 nm strained Si/SiGe p-type MOSFETs [34], based on the calibration with experimental data from [21].

In the third year, Monte Carlo simulations were applied to study the effect of interface roughness on bulk Si and strained Si MOSFETs [35], using a new model

developed within the group [36]. Based on the validation of the model, performance predictions of scaled strained Si MOSFETs were carried out. Soft optical phonon scattering mechanism [6] was also implemented within the Monte Carlo simulator and has been investigated in both conventional Si and strained Si MOSFETs with high-κ gate dielectrics.

1.3 Thesis Outline

A brief description of the content of this thesis is given below:

Chapter 2 briefly reviews current strained Si MOS technology. Different structures, technological issues and solutions of buried channel strained Si/SiGe MODFETs for RF applications and surface channel strained Si MOSFETs for CMOS applications are discussed. Emerging technologies and available therotical work concerning the future of strained Si CMOS technology are also outlined.

Chapter 3 gives a brief introduction of different simulation techniques and discusses the properties of the Si/SiGe heterostructure which are important for numerical device simulations. Calculations and summary of the heterostructure parameters are given. The studied parameters, including the bandgap of strained Si or SiGe, band offsets of the heterostructure at the conduction and valence bands; effective masses; density of states (DOS); and permittivity are essential for all the simulation work associated with the Si/SiGe heterostructure and have been used throughout all the work contained in this thesis.

Chapter 4 presents a comprehensive simulation study of n-type buried strained Si channel MODFETs by using a 1-D Poisson-Schrödinger solver and a 2-D drift-diffusion device simulator MEDICI. The simulations are based on calibrations in respect of a 0.25 μm and a 0.1 μm n-type strained Si/SiGe MODFETs fabricated by DaimlerChrysler. The impact of the device geometry on RF performance and the impact of different device designs on linearity are then investigated.

In Chapter 5, scaling studies for both n-type and p-type bulk Si and strained Si MOSFETs are presented. The simulations are based on calibrations with respect to the sub-100 nm n-type and p-type strained Si MOSFETs fabricated by IBM. Using

drift-diffusion device simulator MEDICI and aiming to CMOS applications, the calibrated n-type and p-type devices are then scaled down to a 35 nm gate length. The impact of parasitic channel in p-type strained Si MOSFETs is investigated and a performance prediction of scaled p-type strained Si MOSFETs is carried out based on hydrodynamic device simulations. The scaled devices are finally used to assess the device and circuit behaviours down to gate length of 35nm.

Using Monte Carlo simulations, Chapter 6 investigates the impact of interface roughness scattering and soft optical phonon scattering on n-type conventional Si and strained Si devices. The interface roughness model developed in the Device Modelling Group has been validated by reproducing the universal mobility behaviour and experimental device characteristics of a 67 nm effective gate length conventional Si MOSFET. The model is then used to evaluate the impact of this scattering on the performance enhancement of strained Si MOSFETs. Based on the validation of the interface roughness scattering model, a performance prediction of strained Si MOSFETs down to a 35 nm gate length is presented. This chapter also contains the simulation work of soft optical phonon scattering induced by high-κ gate dielectrics and compares the impact of such scattering on the device performances between 67 nm conventional Si and strained Si devices with different high-κ gate dielectrics.

In Chapter 7, major contributions of this Ph.D project are summarised. Further work based on existing study is also suggested.

Chapter 2

Strained Si for MOS Technologies -A Literature Review

The lattice mismatch between Si and Ge, which is 4.2% at room temperature, requires either tensile or compressive strain of the active $Si_{1-x}Ge_x$ layer to match the in-plane lattice when it is pseudomorphically grown on a $Si_{1-y}Ge_y$ substrate layer (when $x \neq y$). The strain applied to the active material induces changes in the band structure and forms band offsets between the active material and the substrate material. The details of these strain effects will be discussed in Chapter 3. There are two types of Si/SiGe hetero- interfaces: type I is demonstrated by compressive strained SiGe layer on an unstrained Si substrate, whereas type II is tensile strained Si on relaxed SiGe heterostructure, as illustrated in Fig. 2.1.

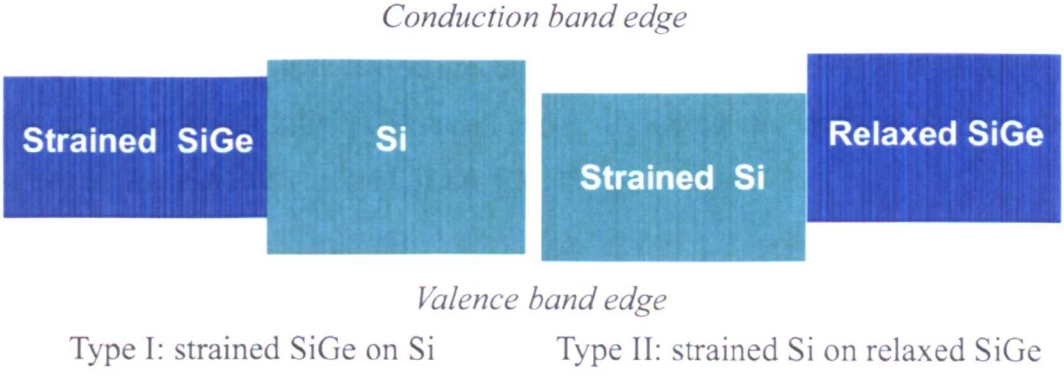


Figure 2.1: The band structures of strained SiGe on unstrained Si and strained Si on relaxed SiGe

The type I (strained SiGe on unstrained Si) heterostructure has a large valence band offset and has been successful in enhancing the performance of p-type strained SiGe channel MOSFETs, as reviewed by Whall and Parker [9,10]. The poor quality of SiO_2 grown directly on the SiGe layer requires a Si cap layer on top of the strained SiGe channel to form an insulator with good interface. However, there is parasitic conduction within the low mobility cap layer. High-κ dielectrics, which may suppress Ge segregation during thermal process, enable the fabrication of a strained SiGe surface channel p-type MOSFET [37 39]. Recent progress has produced a 19 nm ultra-thin body strained SiGe surface channel SOI device with a 2.3 times mobility enhancement over conventional universal mobility behaviour [40]. A 50 nm ultra-thin body fully depleted SiGe channel p-type MOSFET has also been demonstrated with a 70% drive current enhancement [41], although experimental work by Andrieu *et al.* [42] observed less performance enhancement when scaling SiGe channel MOSFETs down to a 50 nm gate length. Nevertheless, the presence of parasitic conduction within the strained SiGe buried channel device and the poor quality of the interface within the strained SiGe surface channel MOSFET are still the major reasons which slows down its applications for CMOS.

This Ph.D project focused on the type II heterostructure, *i.e.*, the strained Si on relaxed SiGe heterostructure. Here, the tensile strain causes band structure changes of Si and as a result enhances both electron and hole transport. Therefore the type II heterostructure has the potential to build both n- and p-type heterostructure field-effect-transistors (HFETs) for performance enhanced CMOS applications. In this chapter, a literature review is made to summarise the (developed and developing) strained Si technologies for future Si based CMOS technology. The focus is on the strained Si for MOSFET technologies, including the strained Si/SiGe for RF applications and the strained Si for CMOS applications.

2.1 Mobility Enhancement in the Strained Si/SiGe Heterostructure

The enhanced mobilities of both electrons and holes are the key features which make this material so promising. When growing Si pseudomorphically on a relaxed SiGe layer, the active Si material is under tensile strain in order to match the lattice constants of the two materials in the plane. As a result of the tensile strain, the lattice constant of tensile strained Si perpendicular to the interface becomes smaller. Fig. 2.2 illustrates the schematics of the Si lattice before (unstrained) and after (tensile strained) growth of the Si layer on relaxed SiGe. The relaxed SiGe on Si substrate therefore behaves like a substrate (buffer) for the strained Si layer and is called a "*virtual substrate*".

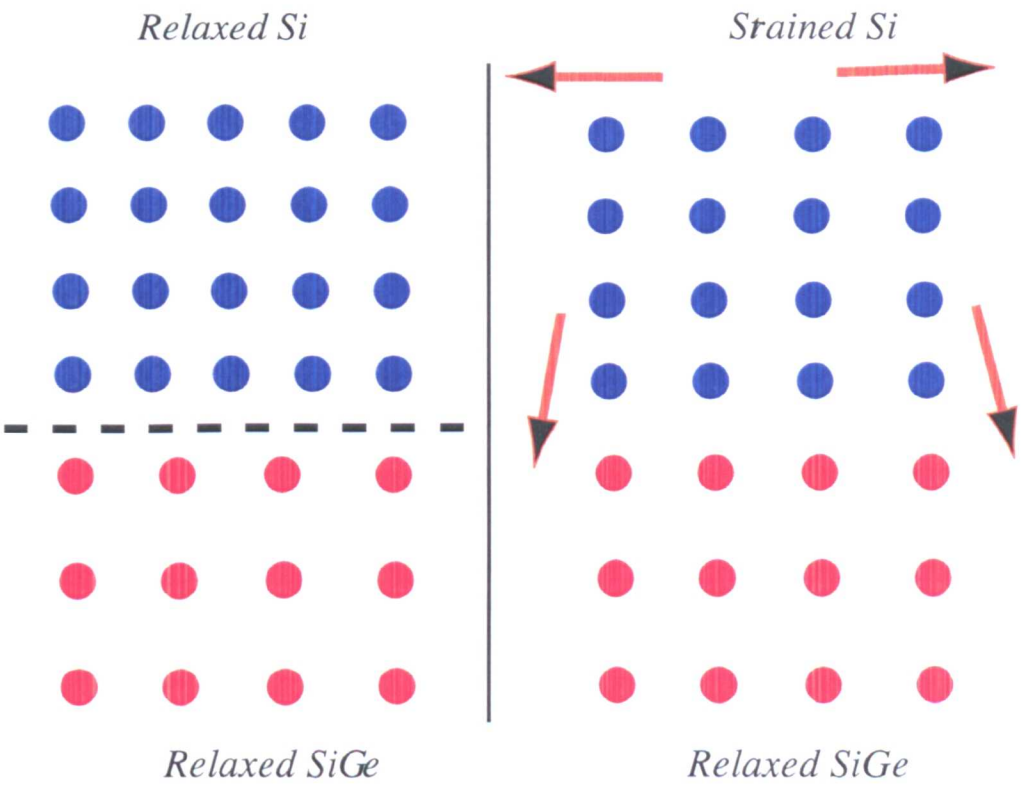


Figure 2.2: Schematics of the Si and SiGe lattice structure. The Si is pseudomorphically grown on the SiGe layer

The strain shifts the energy levels of the valence and conduction bands and splits degenerate bands, which reduces inter-valley scattering, lowers the effective mass, as a result, enhances the carrier transport within the strained layer. Therefore the mobility enhancement of strained Si is dependent on the degree of strain, in other words, the Ge concentration of the relaxed SiGe buffer in this case. Fig. 2.3 and Fig. 2.4 show Monte Carlo simulated low-field electron and hole mobilities [43] in the strained Si layer as a function of Ge concentration within the relaxed SiGe buffer. The electron mobility data has been compared with results from Bufler *et al.* [44].

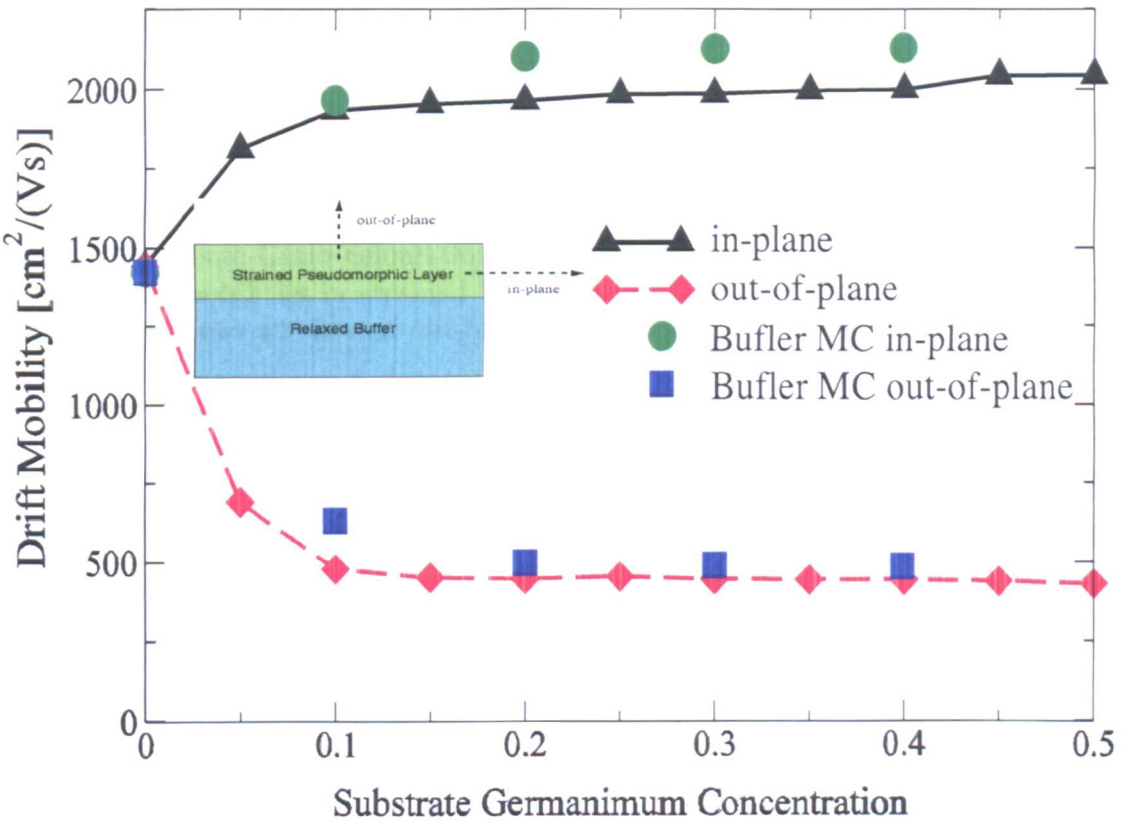


Figure 2.3: Monte-Carlo calculation of low-lateral-field (0.25kV/cm) in- and out-of-plane electron mobilities in strained Si as a function of Ge content within the SiGe buffer. The electron mobilities from Bufler *et al.* [44] are shown for comparison.

From Fig. 2.3, it can be seen that tensile strain enhances the in-plane electron mobility, whereas it degrades the electron mobility in the out-of-plane direction. The in-plane electron mobility enhancement increases as the substrate Ge concentration increases and saturates at a Ge concentration of about 15%-20%.

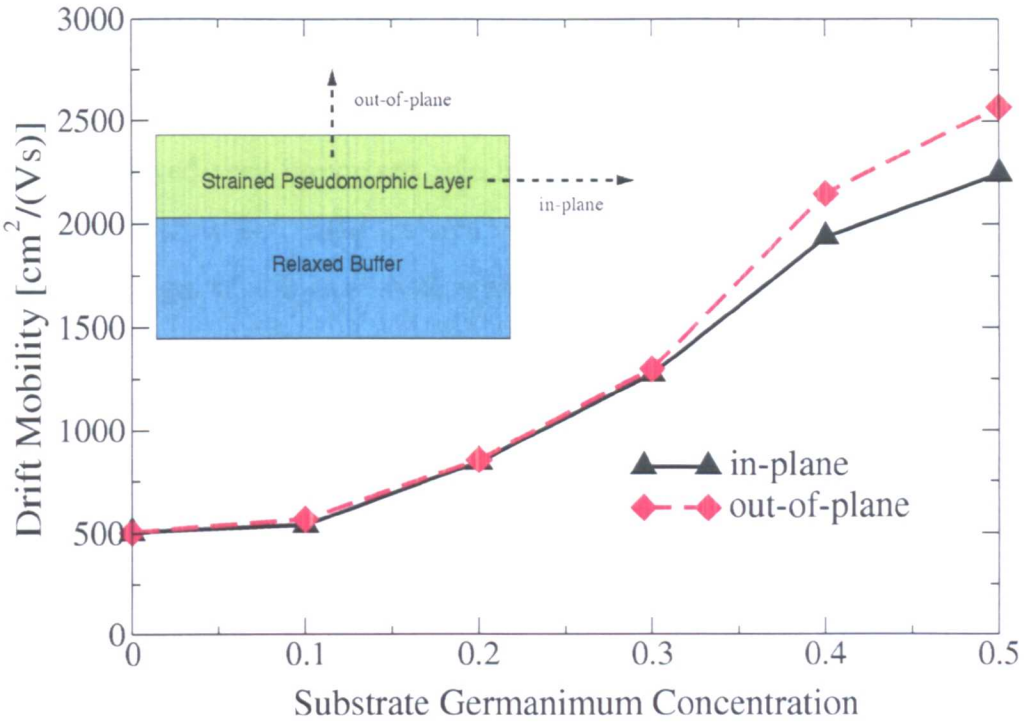


Figure 2.4: Monte-Carlo calculation of low-lateral-field (0.25kV/cm) in- and out-of-plane hole mobilities in strained Si as a function of Ge content within the SiGe buffer. Inset shows the in and out-of-plane directions.

On the other hand, the hole mobility enhancement requires a higher strain, *i.e.*, a large Ge concentration within the SiGe substrate, as indicated by Fig. 2.4. This means that different Ge concentrations of the SiGe substrate are required for p-type and n-type MOSFETs in order to achieve high performance enhancement, which makes the fabrication of CMOS more problematic. Another problem of strained Si for p-MOSFETs is that the degradation of hole mobility enhancement at high electric field is much greater than that of electrons [21]. Such severe hole mobility degradation is due to the fact that high electric field eliminates the benefits from the reduced inter-valley scattering.

Note the substrate orientation studied in this thesis is in the (100) direction. The (110) orientation substrate may be used to enhance the hole mobility because the hole effective mass in the (110) direction is smaller than that in the (100) direction. The hole mobility of a (110) wafer has been reported as more than twice of that of a (100) bulk Si wafer [22, 45]. The (110) direction may solve the performance mismatch problem for CMOS use.

2.2 Various Strained Si Device Structures

Having been studied for more than ten years, Si and SiGe material based band engineering played very important role in improving the performance of Si based CMOS. In earlier stages, SiGe (in compressive strain) on Si heterostructure, having the advantages of enhanced hole mobility and hole confinement, was proposed for p-type MOS applications. However, although the performance for p-type MOSFETs is enhanced, the incorporation of this structure into CMOS isn't promising at present. Recent progress from research groups and major industry vendors indicate that tensile strained Si surface channel MOSFETs, having similar structures and performance enhancements for both n- and p-type devices, is a better choice for performance enhanced CMOS applications. Without significant additional cost, the strained Si MOSFET may be realised via different techniques and has demonstrated remarkable performance compared to bulk Si CMOS, showing a bright future for CMOS technology.

In the area of RF applications, although III-V remains the pacemaker even today, designs based on SiGe technology have been developed, such as the SiGe HBT and the Si/SiGe heterostructure FETs (HFETs) for RF applications.

2.2.1 Strained Si/SiGe Modulation Doped Field Effect Transistors (MODFETs)

The type-I Si/SiGe heterostructure (see Fig. 2.1) is favourable for hole confinement and has been exploited in many heterostructure devices, *e.g.* the p-type MOSFETs/MODFETs and the SiGe HBT [14,46 50]. The type II structure (see Fig. 2.1) with high electron and hole mobilities is potentially better than the type I structure for CMOS applications. The focus of this project is the n-type strained Si/SiGe HFETs, mainly modulation doped field effect transistors (MODFETs) [46 53].

The basic structures of strained Si channel MODFETs are shown in Fig. 2.5. The multi-layer MODFETs have undoped strained Si channels with side doped SiGe supply layers. Spacer layers are added between the channel and the supply layers to reduce the effect of ionized impurity scattering on the mobility of channel. Car-

riers are supplied from remote doped SiGe layers, a process which is known as a *modulation doping*.

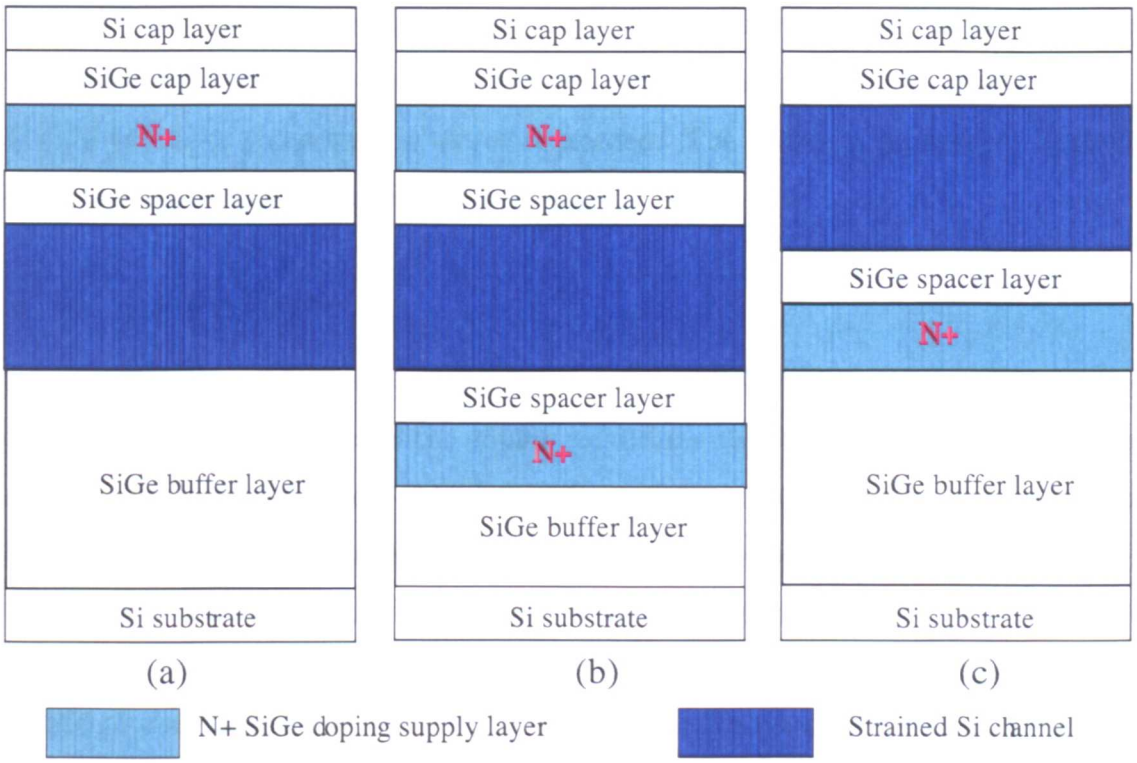


Figure 2.5: (a) Front-, (b) double- and (c) back- side modulation doped strained-Si/SiGe MODFETs

Like conventional Si MOSFETs, Si/SiGe MODFETs also have two working modes: the depletion and enhancement modes. The enhancement mode in a MODFET with Schottky gate may be achieved by gate recessing with varying recess depth [48, 51]. In the enhancement mode MODFET, the gate is located in, or close to, the undoped spacer layer. At zero gate bias, the channel beneath the gate is entirely depleted. Therefore, the gate has to be forward-biased for transistor action. The adjustment of the operation mode can also be made by changing the doping in the supply layers and the thicknesses of layers above the channel. Recent research has concentrated mainly on the depletion mode MODFET. Utilizing the structure show in Fig. 2.5(b), a 0.1 μm n-type $Si/Si_{0.6}Ge_{0.4}$ MODFET with f_T—74 GHz at 300 K [54] and a 0.1 μm n-type $Si/Si_{0.6}Ge_{0.4}$ MODFET with f_{max}—188 GHz at 300 K and f_{max}—230 GHz at 50 K have been successfully demonstrated [25]. Because of the suppression of noise

sources from the interface the structure may also be used for low noise applications, as has been demonstrated by a 0.3 dB minimum noise figure at 25 GHz in a 0.13 μm $Si/Si_{0.58}Ge_{0.42}$ n-MODFET [26].

To achieve high transconductance that is required for RF applications, the high mobility carrier density in the channel needs to be maximized [30,55]. The optimization of the device structure is therefore needed. The detailed impacts of different layers and gate structures are outlined below.

2.2.1.1 Layer Issues

As shown in Fig. 2.5, there are different possible MODFET layer structures. The thicknesses of these layers and the doping conditions need to be optimized in order to obtain a high concentration of high mobility carriers confined within the quantum well [30, 55]. The thin undoped Si cap layer is used to protect against the in-depth oxidation into the device and can reduce the effects from interface roughness scattering. Changing the cap thickness also adjusts the control of the gate over the 2DEG (2-dimensional electron gas) in the channel. Modulation doped layers supply high mobility carriers to the channel through the spacer layers. Increasing the spacer layer thickness reduces the Coulombic scattering from the ionized impurities in the doped layer. However, the carrier density in the channel decreases when the spacer becomes too thick. Higher doping concentrations in the supply layer increases the carrier density within the channel. Again care has to be taken as a parasitic channel may be created in the layers above the channel. Careful adjustment of the doping levels and the layer thicknesses can reduce the parasitic current to an acceptable level.

Modulation doped layers may be located above the channel, or below the channel or both (above and below), as illustrated in Fig. 2.5. The drawback of the front-side (above the channel) doping structure is the early onset of the parasitic channel above the channel, which is not observed in those structures with backside (below the channel) doping. Too high a backside doping may also lead to the formation of a parasitic current below the channel that does not allow the device to pinch-off completely [53]. The carrier density in the double-side doped channel structure

(Fig. 2.5b) is nearly doubled which is useful for those applications requiring a low channel resistance. This type of structure permits a symmetrical wave function in the rectangular channel which maximizes the carrier density in the channel [48].

2.2.1.2 Buffer Issues

From the design point of view, the most important problems are the optimization of the cap thickness, the doping concentration and the spacer thickness. However, in real devices, the SiGe virtual substrate is also a key factor affecting the device performance. The difficulty lies in realizing a thin high quality SiGe buffer. Different types of buffers have been proposed, such as the stepwise Ge content buffer and the graded Ge content buffer [56 60].

Good SiGe buffers need to be fully relaxed and act as the buffer between the Si substrate and the strained Si layer. However, a high density of threading dislocations may be created in the buffer during its growth. The growth temperature and the growth Ge gradient are the key growth parameters. Low growth temperatures lead to shorter misfit dislocation lines and a correspondingly higher threading dislocation density [59]. A low grading rate, *i.e.*, a thick buffer, has fewer defects [58], but too thick a buffer costs much more in terms of growing time, making it less compatible with existing Si processes. Thick buffer also induces self-heating problem due to the much lower thermal conductivity of SiGe [61]. Ref. [46] suggests a Ge content grading rate of between 20%/μm and 30%/μm to be optimal. A higher defect density not only affects device performance, but also limits the commercial applications of this technology. To deal with this problem, there have been many proposed approaches to grow buffer. Hackbarth *et al.* [62] recently studied different approaches to grow micron thickness $Si_{0.7}Ge_{0.3}$ buffers and concludes that overgrown graded UHVCVD (ultra high vacuum chemical vapour deposition) and LEPECVD (low energy plasma enhanced chemical vapour deposition) buffers have great advantages over entirely MBE (molecular bean epitaxy)-grown samples. Another particularly interesting technique is to grow the SiGe buffer on an SOI (Si/SiO_2) substrate using the separation by implanted oxygen (SIMOX) technique [63,64]. This type of structure allows the relaxation of a thin SiGe buffer layer with the dislocations threading

down into the Si/SiO_2 substrate and is potentially one of the best ways to provide a thin SiGe buffer good enough for strained Si applications. Another advantage of using a SOI based buffer is to reduce junction capacitances and suppress leakage current from the buffer, thereby improving device performance.

2.2.1.3 Gate issues

The differences between two types of gate versions, the Schottky-gate and MOS-gate HFETs, have been investigated in [65, 66]. The reported highest cut-off frequency (f_T) of 72 GHz in a 0.1 μm n-MODFET [54] and maximum oscillation frequency (f_{max}) of 188 GHz in a 0.1 μm n-MODFET [25], are all achieved using Schottky gates. Although the Schottky-gate HFETs have a higher frequency performance than that of MOS-gated devices, they suffer from a reduced gate voltage swing due to the onset of a significant gate leakage current at high gate bias. Ref. [66] reported that the gate leakage current of Schottky gate devices is about 3 orders of magnitude higher than that of MOS-gated devices.

The gate structure used in MODFETs has a larger source-to-drain distance than the gate length. This may be useful to reduce short-channel effects, such as drain-induced-barrier-lowering (DIBL), and reduce the parasitic capacitance caused by dopant diffusion under the gate from the source and drain [65]. The disadvantages are an increased source-drain intrinsic resistance and a lower transconductance [65].

2.2.2 Strained Si MOSFETs

The advantages of strained Si for CMOS include the enhancement of mobility for both electrons and holes and the compatibility with existing Si technology. Achieved either by the widely used technique to grow strained Si on a SiGe virtual substrate or by recently emerged process stress techniques, strained Si technology has been accepted as the most promising solution for the enhancement of near future CMOS in the roadmap [1].

2.2.2.1 SiGe Virtual Substrate Based

The research of strained Si for MOSFETs has been mainly focused on relaxed SiGe *virtual substrates* since the strained Si n- and p- MOSFETs on relaxed SiGe buffer were first demonstrated in 1992 [15]. The surface channel strained Si MOSFET structure is similar to that of the conventional bulk Si (or *control* Si) MOSFET. Some additional process considerations are required for the fabrication of strained Si MOSFETs, however these are not the major interest of this work.

A well-known disadvantage of strained Si for CMOS applications is that the hole mobility enhancement in strained Si is much less than that for electrons. Indeed, for a long time, most of the reported performance enhancements in strained Si MOSFETs were for n-type devices.

In 1992, Welser *et al.* [15] first demonstrated long channel n- and p-type strained Si on relaxed $Si_{0.7}Ge_{0.3}$ MOSFETs, showing a 2.2 times larger electron mobility in the n-type strained Si MOSFET but no difference in the p-type strained Si MOSFET compared to *control* Si MOSFETs. Two years later, Welser *et al.* [67] measured the strain dependence of electron mobility enhancement, demonstrating that the enhancement factor of the mobility in the strained Si layer over that in the Si saturates at 1.76 when the Ge content was increased to about 20% within the SiGe *virtual substrate.*

Rim *et al.* [68] first reported the strain dependence of the hole mobility enhancement in long channel p-type strained Si on relaxed SiGe MOSFETs. It was found that increasing the substrate Ge content enhances the hole mobility and at 29% Ge content the device shows a 1.8 times larger hole mobility than that of the bulk Si device.

However, due to high electric field and high channel doping, these low-field mobility enhancements degrade in small devices. The high channel doping in short channel MOSFET helps to suppress the SCE effects, but on the other hand it induces high electric fields and causes degradation of the mobility. Rim *et al.* [69] demonstrated a 0.1 μm n-type strained $Si/Si_{0.8}Ge_{0.2}$ MOSFET, showing a 45% improvement in transconductance and a 75% electron mobility enhancement at high electric fields. They also reported the self-heating effect in the thick SiGe layer,

which requires a AC measurement to avoid this effect. In a DC measurement, the self-heating effect degrades the drain current by 10-20% [61].

In 2001, Rim *et al.* [31] reported a 67 nm effective gate length n-type strained $Si/Si_{0.85}Ge_{0.15}$ MOSFET with a 70% increase in electron mobility at high electric field (1.5MV/cm) and a 35% drive current increase. The mobility enhancement at such high electric field suggests a new mobility enhancement mechanism [31], which according to Fischetti *et al.* [70] and the conclusion in Chapter 6 of this thesis is an "*improved*" interface roughness of the strained Si/SiO_2 interface, *i.e.*, reduced interface roughness scattering in strained Si MOSFETs. Based on this breakthrough, IBM then announced their plans to use strained Si for their future CMOS technology and reported a 35% performance boost with only a 10% cost increase based on existing Si CMOS technology [16]. Later, a sub-100 nm p-type strained $Si/Si_{0.72}Ge_{0.28}$ MOSFET was demonstrated by Rim *et al.* with a 45% enhancement of peak hole mobility and 7-10% drive current increase, as well as a n-type strained $Si/Si_{0.8}Ge_{0.2}$ MOSFET with a 110% enhancement at high electric field (1.5MV/cm) and >15% drive current increase [21, 71].

Recent progresses by AMD include the fabrication of n-type strained $Si/Si_{0.8}Ge_{0.2}$ MOSFETs down to 25nm gate length [23] and the demonstration of a 45% drive current enhancement in a 35nm physical gate length NiSi metal gated device.

In terms of circuit performance, in a 70 nm strained Si CMOS technology compared to bulk Si CMOS, an 86% electron mobility enhancement and over 20% increase in the drive current have been demonstrated by Huang *et al.* [17], delivering a 95% higher inverter peak current and a reduction of 2.2 ps in delay of the ring oscillator at the same drive current. Wang *et al.* [24] recently reported a 60 nm gate length strained Si CMOS with the lowest ring oscillator delay of 6.5 ps at 1.2 V operation.

However, there are some very important fabrication issues which need to be considered for strained Si CMOS [72], including the misfit dislocations associated with the SiGe buffer, and thermal budget constraints. The misfit dislocations are induced by relaxation processes of the SiGe buffer and, at the strained Si/SiGe interface, increase the diffusion coefficient of Arsenic by up to six orders of magnitude [73]. More

experiments on strained $Si/Si_{0.8}Ge_{0.2}$ have confirmed the increased diffusivities of n-type dopants in relaxed SiGe [72, 74]. In contrast to n-type dopant diffusion, boron diffusion was found to be suppressed with increasing Ge fraction [72, 75]. These diffusivity changes therefore necessitate a greater control of doping. The thermal processing, on the other hand, will affect dopant diffusion, strain relaxation and interdiffusion at the Si/SiGe interface. Koester *et al.* [76] studied the effect of thermal annealing on strained Si on SiGe heterostructures by using Raman spectroscopy. They found that annealing at 1000°C for as long as 300 seconds resulted in interdiffusion at the Si/SiGe interface and therefore effectively reduced the strained Si thickness, but didn't cause significant strain relaxation of strained Si layer.

Strain relaxation, which leads to the introduction of misfit dislocations, may start to occur once the thickness of the strained Si layer exceeds the *critical thickness.* Recent experimental work by Fiorenza *et al.* [77] demonstrated a low field electron mobility enhancement for strained Si on $Si_{0.8}Ge_{0.2}$ of about 80% on all samples, with strained Si layer thickness both less and far greater than the critical thickness (14nm). The explanation of the observed mobility enhancement within a strained Si layer thicker than the *critical thickness* is that carrier conduction is occurring far away from the misfit dislocations induced by partially relaxation. However, the increased misfit dislocation density at the Si/SiGe interface in the sample with thicker strained Si layer forms dopant diffusion pipes between source and drain, leading to an increase in off-current. It is therefore concluded [77] that the chief motivation to limit the strained Si thickness below the *critical thickness* is to minimize the leakage current, rather than minimize the loss of strain and mobility enhancement.

2.2.2.2 Process Straining

Process strained Si may be achieved through the proper strain engineering of the shallow trench isolation (STI), the cap layer (contact-etch-stop layer) or silicidation processes [11 13]. These three process induced strain techniques have different effects on the drive current of MOSFET, as concluded by Ge *et al.* [12]. Ge *et al.* also studied the 3D strain sensitivity of CMOS performance enhancement and found that the application of strain (tensile or compressive) in different directions (x or

y or z) has different effects on the drive current of n- or p-type MOSFETs. Each approach results in a 5-10% enhancement of ring oscillator speed [12]. Using combinations of these strain techniques, the 3D process induced strain is able to enhance the drive currents of both n- and p-type MOSFETs. The strains were also found to have different dependences on the channel length and channel width [12, 13].

Intel has announced its strained Si applications within its current 90nm technology [11]. Intel's strained Si CMOS uses a high stress Si_3N_4 cap layer to achieve tensile strain in the channel of an n-MOSFET and a SiGe source-drain to obtain compressive strain in the channel of a p-MOSFET device, which deliver increases in drive current of 10% for n-MOSFET and 25% for p-MOSFET with only an extra 2% cost based on existing Si CMOS technology [11].

Chan *et al.* [13] reported IBM's high speed 45 nm gate length strained Si MOSFETs, which was realized by the integration of their 90 nm bulk Si technology with a strain engineering. These devices, optimized by stress effects from both STI and contact etch stop nitride films, showed an 8% increase in drive current in n-MOSFET and no degradation in p-MOSFET and therefore improved the ring oscillator speed by 5%.

2.2.3 Strained Si SOI MOSFETs

Takagi *et al.* [78] summarized the current issues of strained Si MOSFETs and the advantages of strained Si SOI MOSFETs. Possible problems of designing strained Si on relaxed SiGe MOSFETs [72, 78] include: (1) Mobility degradation due to Ge outdiffusion, which is associated with the strained Si thickness and thermal budget during fabrication; (2) The higher Ge content of the SiGe substrate required for a notable hole mobility enhancement in p-MOSFET, which on the other hand worsens the quality of the strained Si layer and the SiGe buffer; (3) Process control for the different diffusion rates of boron, arsenic and phosphorus in SiGe; (4) The strain induced threshold voltage shift and junction leakage due to band offsets of the heterostructure and the narrowed bandgap of strained Si; (5) The self heating of a thick low thermal conductivity SiGe buffer.

On the other hand, the strained Si SOI structure has inherent immunity to

some of the problems mentioned above, in which the relaxed SiGe layer on insulator (usually called SGOI) acts as the *virtual substrate* of the strained Si layer. With a thin SiGe buffer, especially in a fully depleted (FD) SOI device [4, 78], (1) the mobility is enhanced because of much reduced channel doping; (2) the junction capacitance and junction leakage are reduced; (3) the self-heating effect due to a thick SiGe buffer diminishes; (4) the SCE effects are suppressed resulting in the much improved scalability of this type of structure.

Many efforts have been made to improve the quality of the SGOI wafer and reduce the buffer thickness. Recent progress in realizing strained Si directly on insulator (SSDOI or SiGe Free) has been reported [12,20,79]. The details are discussed below.

2.2.3.1 SiGe Virtual Substrate Based

Different technologies, such as: SIMOX [63, 64, 80–82]; wafer bonding and layer transfer [83] and Ge condensation (enrichment) techniques [84–88], have been investigated to grow the SGOI structure. SIMOX technology is, perhaps, the most mature technology used to grow the SGOI structure. However, the temperature of the SIMOX annealing process is higher than the melting temperature of SiGe material when the Ge content is >10% [4,87,88]. Such low Ge contents are not sufficient to supply enough strain to achieve hole mobility enhancement [89, 90].

The Ge condensation (enrichment) technique, a combination of the internal-thermal oxidation (ITOX) and SIMOX techniques, has been used to grow ultra-thin high Ge content SGOI substrates [84,88]. Using this technique, a long channel fully depleted SOI MOSFET based on a 7 nm thick strained Si layer on a 53 nm thick 25% Ge content relaxed SiGe layer has been recently reported [88], showing 85% and 53% maximum enhancements in electron and hole mobility, respectively. Ring-oscillators constructed from these high-speed strained Si SOI transistors demonstrated a 63% improvement in speed over the control-SOI CMOS device at the a supply voltage of 1.5 V. The technique also enables the deposition of ultra thin (9 nm) high Ge content (56%) SiGe buffers [84]. High uniformity and low dislocation density ($<10^8$ m^{-2}) of films were reported using this technique [78]. Lee *et al.* [91] reported their

sub-70 nm strained Si on ultra-thin (T_{SOI}<55 nm) thermally mixed SGOI substrate technology. These devices are based on IBM's conventional SOI technology and the overall performance enhancement of n-MOSFET is 20-25% over the conventional Si SOI MOSFET. In these devices, the interface roughness RMS (root mean square) height is reported to be 0.52 nm.

The wafer bonding and layer transfer approach has the advantage of growing highly relaxed high Ge content SiGe layers, but the technique has the difficulties of growing thin layer and controlling the quality and uniformity of the SiGe layer [4]. Using this technique, long channel devices have delivered a 50% electron mobility enhancement (with 15% Ge content of SiGe buffer) and a 15-20% hole mobility enhancement (with 20-25% Ge content of SiGe buffer) [92]. In these devices, the reported RMS height of bonded SGOI wafer after CMP (chemical mechanical polishing) is 0.4 nm.

2.2.3.2 SiGe-Free Strained Si Directly on Insulator

Although progress has been made towards thinning the SiGe buffer layer, the large thickness of the buffer still limits the scaling of devices and the implementation of FD SOI below the 35 nm regime. Indeed, the problems associated with the presence of a SiGe layer, like the changed diffusion rates of dopants in the SiGe layer and the diffusion of Ge into strained Si layer, still exist. The SiGe free or strained Si directly on insulator (SSDOI) structure was therefore proposed to avoid the obstacles from SiGe layer but maintain the mobility enhancement possible from the strained Si channel.

Lando *et al.* [19,93] first reported their SiGe free buffer for strained Si MOSFETs by wafer bonding and hydrogen-induced layer transfer. A 49 nm strained Si layer with RMS of 9.5 Å was transfered after its growth on a relaxed $Si_{0.68}Ge_{0.32}$ buffer onto a 100nm SiO_2 substrate.

Rim *et al.* [20] first demonstrated their long channel SSDOI MOSFETs with performance enhancements for both n- and p-type devices. IBM intends to use these devices in its future CMOS technology. Ultra thin (<20 nm) strained Si layers with RMS of <2 nm were transfered from a high Ge content SiGe layer (>35%) to a 120

nm SiO_2 substrate. Long channel devices with an inversion carrier density of 10^{13} cm^{-2} were demonstrated with an electron mobility enhancement of 125% and a hole mobility enhancement of 21%. Sub-60 nm n- and p-type SSDOI devices with good subthreshold characteristics have also been reported, while further optimizations are required to suppress the series resistances to optimise the performance enhancement.

Yin *et al.* [79] reported their strained Si on insulator without SiGe buffer technology. A strained Si layer (approximately 20 nm thick) with an equivalent strain on a relaxed $Si_{0.85}Ge_{0.15}$ buffer was transfered to a compliant borophosphorosilicate glass (BPSG) insulator by wafer bonding and 'Smartcut' processes. A 55-60% higher transconductance and effective electron mobility near turn-on were shown in a long channel n-type strained Si on insulator device over a bulk Si SOI MOSFET.

Recent work by Lauer *et al.* [94] reported that the mobility enhancement in Ge-free strained Si on insulator n-MOSFETs, with strained Si thickness far greater than the *critical thickness,* may be obtained. This fact is further supported by the same behaviour exhibited in bulk strained Si/SiGe MOSFETs [77], which was mentioned in last section. Lauer *et al.* also showed an immunity to misfit dislocation at the Si/substrate insulator interface, which in fact solves the leakage current problem which was reported in strained Si on relaxed SiGe MOSFETs [77]. The results indicate that a possibility of growing strain Si layer with a thickness thicker than the previously understood *critical thickness.*

In general, the SSDOI technology is very promising for future ultra scaled FD SOI MOSFETs, in which SCE effects are intrinsically controlled and the mobility is enhanced due to strain in the channel, low vertical electric fields and low channel doping. The structure also possess immunity to the leakage current induced by the misfit dislocation at the Si/SiGe interface. Using this technology, it is therefore possible to benefit from the performance and scalability of the strained Si double-gate MOSFET (without the SiGe buffer), although uncertainties over fabrication still exist [4].

2.2.4 (110)-orientation Strained Si MOSFETs

Besides the widely used (100) orientation wafer for the MOSFETs discussed above, the (110) orientation has also recently been proposed for CMOS applications, demonstrating remarkable hole mobility enhancements [22, 45]. The smaller hole effective mass along the (110) orientation along with reduced inter-valley scattering result in an appreciable hole mobility enhancement for strained Si in the (110) orientation. It may be used in the p-MOSFET in CMOS, balancing the performance between n-type and p-type devices.

Using the Ge condensation technique, n- and p-type strained Si SOI MOSFETs have been fabricated by Mizuno *et al.* [45] on a (110) relaxed SGOI substrate with a Ge content of 25%. The resulting devices show electron and hole mobility enhancements, of 81% and 203% respectively, over the universal mobilities of (100) bulk MOSFETs.

Yang *et al.* [22] recently reported IBM's hybrid-orientation technology (HOT), which implements a (100) Si n-MOSFET and a (110) Si p-MOSFET on the same wafer to achieve high performance CMOS. This technology uses wafer bonding and layer transfer techniques and has demonstrated a substantial drive current enhancement for n-MOSFET. An 80 nm gate length (110) Si p-MOSFET delivered a 65% higher drive current than a conventional (100) bulk Si p-MOSFET even with a non-optimized surface roughness within the (110) device. This allows the performances of n- and p-MOSFETs in CMOS applications to become balanced. Based on the current strained Si and HOT technologies, although the fabrication might be challenging, future high performance CMOS may be achieved by integrating strained Si channel, SOI and different orientations together on a wafer.

2.2.5 Strained Si MOSFETs with High-κ Dielectrics

The continuous scaling of MOSFETs driven by the need for increased performance beyond the 45 nm technology node requires extremely thin gate oxides, resulting in intolerably high gate leakage and oxide/interface uniformity problems during device fabrication [1]. High-κ dielectrics may then be used to replace the long studied

SiO_2 to enable further scaling. Although high-κ dielectrics significantly reduce the leakage current and improve scalability, the qualities of the Si/insulator interface and the dielectrics itself are currently not appropriate for commercialization.

Rim *et al.* [95] first demonstrated a sub-100 nm strained Si MOSFET with an inversion equivalent oxide thickness (EOT) HfO_2 of 2.8 nm, showing a 1000 times reduction in gate leakage compared to a conventional Si MOSFET with an an EOT SiO_2 of 3.1 nm. Although a notable electron mobility enhancement of strained Si with HfO_2 over bulk Si with HfO_2 or SiO_2 has been observed, the experimental results also show that the mobility in an HfO_2 based MOSFET (either strained Si or bulk Si) is much degraded from that of corresponding SiO_2 based MOSFET. This analysis indicates that a new mobility limiting mechanism is required to explain the mobility degradation associated with high-κ dielectrics. This degradation may be due to Coulomb scattering from the trapped and fixed charges [95] or soft-optical phonon scattering [6]. Recent experimental studies by Ren *et al.* [7] and Datta *et al.* [96] suggest that soft-optical phonon scattering is a intrinsic limiting factor for high-κ induced mobility degradation and cannot be counteracted by process optimization.

Datta *et al.* [96] reported sub-100 nm n-type strained Si channel MOSFETs with an HfO_2 dielectric of EOT=1 nm and a TiN metal gate electrode, showing a 35% mobility enhancement at 1 MV/cm effective field and a 1000 times reduction in gate leakage current compared to unstrained Si devices with conventional oxides. It was also concluded that a metal gate with a higher free electron concentration and a higher plasma frequency may dynamically screen the high-κ soft-optical phonons by coupling with the inversion carriers and is therefore beneficial in improving the soft-optical phonon limited mobility [96].

Recent work by Zhu *et al.* [97, 98] on mobility measurements for HfO_2-gated MOSFETs confirmed that both Coulomb scattering and soft-optical phonon scattering play important roles in the mobility degradation due to the introduction of high-κ dielectrics.

2.3 The Future of Strained Si - How Far It Can Go?

Strained Si has demonstrated high performance and compatibility with existing Si technology. That this material may replace Si channels has been widely accepted by the industry and has been considered as the first step towards the future of enhanced Si based CMOS technology, as indicated in the latest version of the International Technology Roadmap for Semiconductors [1]. The introduction of high-κ dielectrics, which is scheduled to be a necessary step beyond the 65 nm technology node to solve the gate leakage problem and enable further scaling, is still problematic and requires further investigations in order to be better understood [1]. In contrast to the status of high-κ dielectrics, strained Si is ready to be integrated into existing Si technology, an example of which is the Intel's 90 nm strained Si technology announced in 2003 [11].

When the gate length is reduced to 10-20 nm, the ultra-thin (UT) body fully-depleted (FD) SOI or double-gate MOSFET (DG-MOSFET) structure is required to have good control of short channel effects and high performance [1, 4]. However, regardless of the fabrication issues, the aggressive scaling of UT FD SOI or DG-MOSFETs is intrinsically limited by the mobility degradation observed in the ultrathin channel. This is due to increased phonon scattering [99, 100], long-range (remote) Coulomb scattering (RCS) induced by the heavily doped source/drain [101] and polycrystalline gate [102 105], surface roughness scattering contributed by both top and bottom Si/insulator interface [106], and possibly soft optical phonon scattering [107].

The phonon limited mobility in the UT SOI device is nearly independent of the channel thickness, t_{channel}, when it is greater than approximately 5 nm [108]. As the thickness is reduced, the phonon limited mobility starts to increase, reaching a maximum at a critical thickness, t_{c}, which is ~3 nm from [108, 109] and ~3.5 nm from [107]. When t_{channel} is less than t_{c}, the phonon scattering rate increases dramatically and the phonon limited mobility decreases monotonically. When t_{channel} is less than 10nm, the presence of the buried (substrate) Si/SiO_2 interface within the UT SOI modifies the surface roughness scattering induced by the gate Si/SiO_2 interface and provides a non-negligible scattering rate to the total surface roughness scattering [106]. The total surface roughness kills the phonon-limited mobility increase

observed when $t_{channel}$ is between 3nm and 5nm as mentioned above and the total effect of the surface roughness and phonon scattering forces a sharp degradation in the mobility if $t_{channel}$<3 nm [109]. Long-range Coulomb scattering, which doesn't affect carrier transport in large devices, is non-negligible in small devices and affects inversion layer mobility via the interactions between the inversion layer electrons and the electrons in the heavily doped source/drain or gate. By self-consistent full-band Monte Carlo/Poisson simulations, Fischetti *et al.* [101] observed as much as a 50% degradation in the effective electron velocity due to remote Coulomb scattering for devices with gate lengths shorter than about 40nm and oxides thinner than 2.5-3 nm, and a 20-25% mobility decrease at a large sheet carrier density for an oxide thickness of 1.5 nm [102]. Saito *et al.* [103] theoretically studied remote Coulomb scattering and reported a 20% peak mobility reduction within a 1-nm-thick SiO_2 MOSFET. Gámiz *et al.* [104] suggested that remote Coulomb scattering cannot be neglected for oxide thicknesses below 2nm, but the effect is negligible for oxide thickness above 5 nm. Esseni and Abramo [105] calculated the electron mobility based on the relaxation time approximation and suggested that the remote Coulomb scattering induced mobility degradation should be clearly observed for oxide thickness below approximately 3 nm. The work by Esseni and Abramo also suggested that high-κ gate dielectrics may help to suppress the remote Coulomb scattering, although this would introduce additional issues, such as fixed charges, interface states and soft optical phonon scattering.

The introduction of a back-gate within DG SOI MOSFETs enables better gate control over the channel and modifies the scattering mechanisms discussed above. The DG SOI MOSFETs [109] operate in *volume inversion* when the channel thickness is less than the sum of the depletion regions induced by the two gates and enables carrier distributions throughout the entire channel. The presence of *volume inversion* reduces the effect of all scattering mechanisms [109] and enables higher mobility than that in single-gate (SG) SOI MOSFETs for devices with channel thickness between 5 nm and 20 nm. However, when considering all the scattering mechanisms mentioned above, for $t_{channel}$<5 nm, both the SG SOI and DG SOI devices exhibit a dramatic mobility degradation which therefore limits the further

scaling of these kinds of device structures [109].

Another important requirement for further device scaling is the application of high-κ gate dielectrics. However, regardless of the fabrication issues [5], soft optical phonon scattering (induced by the strong coupling between carriers in the inversion layer and low-energy surface optical phonons arising due to the highly polarized bonds within high-κ dielectrics [6]) and remote Coulomb scattering (due to the existence of fix charges and interface states) strongly affect carrier mobility in the channel and are intrinsic limitations of device performance. The use of a transport enhanced material in the channel, such as strained Si or Ge, is therefore necessary in order to maintain a relatively high channel mobility in MOSFETs with high-κ gate dielectrics.

However, although strained Si channels enhance the carrier transport within the channel, and partially counteract the mobility degradation of extremely scaled thin body devices, the strained Si channel thickness of either the SG SOI or DG SOI MOSFET is still limited by the mobility degradation within an extremely scaled channel [110 112]. The further scaling of either conventional Si or strained Si MOSFETs is therefore limited by the channel thickness at ~3nm.

Aside from strained Si channel devices, some other promising device structures based on Si/SiGe heterostructure are also under development. Dual channel MOSFETs have shown to be very interesting for CMOS applications. Tensile strained Si on compressively strained $Si_{0.2}Ge_{0.8}$ on a relaxed $Si_{0.5}Ge_{0.5}$ substrate MOSFET, in which the strained Si is used for electron transport and the strained SiGe for hole transport, was demonstrated by Leitz *et al.* [113], delivering a hole mobility enhancement factor of 5.15 over a conventional Si MOSFET. A recent promising work for CMOS applications, the strained Si on strained Ge dual channel MOSFET by Lee *et al.* [114], has been reported with hole and electron mobility enhancement factors of 10 and 1.8 times, respectively, over conventional Si MOSFETs. Furthermore, high-κ gate dielectrics also enable a high mobility Ge layer to be used as a channel within a MOSFET, as demonstrated by Chui *et al.* [115] for an n-MOSFET and by Ritenour *et al.* [116] for a p-MOSFET.

2.4 Summary

This chapter has reviewed the current status of strained Si based technology. The devices discussed are based on the strained Si/SiGe heterostructure and have been classified into two categories here: buried channel strained Si channel MODFETs for RF applications and surface channel strained Si MOSFETs for CMOS applications. The strained Si/SiGe MODFETs have multi layers and a T-shaped Schottky gate, which require careful optimization for specific high performance applications. Strained Si MOSFETs, which range from bulk structures to single gate SOI and double gate SOI, have been discussed in detail by describing device structures, associated advantages and disadvantages, technology solutions and current status. The application of high-κ dielectrics is also discussed as one of the solutions for the further scaling of CMOS. With regard to simulations, theoretical work on scattering mechanisms associated with ultra scaled devices have been outlined to discuss the limit of device scaling.

Based on this literature review, it is clear that simulation work is necessary to optimize device design and to gain insight into device operation. As a basis, a detailed understanding of the Si/SiGe heterostructure is therefore required. This is addressed in next chapter.

Chapter 3

Simulation Techniques and Si/SiGe Heterostructure Parameters

Numerical device modelling is playing an ever increasing important role in the semiconductor industry. By using device simulation it is possible to assess device behavior before fabrication. Various commercial numerical device simulators are available, such as MEDICI and TAURUS from Synopsys [27], ATLAS from Silvaco [117] and DESSIS from ISE [118]. These commercial tools are often used along with other non-commercial simulators, such as Poisson-Schrödinger solvers and Monte Carlo simulators. The selection of these tools necessitates an understanding of the validity and limitation of the physical models and numerical methods used in different simulators. The 2-D device simulators, MEDICI and TAURUS, employing Drift-Diffusion (DD) and Hydrodynamic (HD) models, together with the ensemble Monte Carlo simulator developed in the Device Modelling Group etc, were used in this Ph.D project. These different models solve the Boltzmann transport model at different levels of approximation and thus may be used for different purposes. Apart from understanding the physical models employed within those simulators it is also essential to have knowledge of the parameters associated with the device simulation, particularly when studying devices constructed from new materials. To date, there has been some confusion regarding the parameters for advanced materials, strained Si or SiGe in this case, and a complete set of associated parameters are required to replace the default parameters provided by the commercial TCAD tools. For

the purpose of this Ph.D project a study of the Si/SiGe heterostructure parameters becomes essential before simulating any strained Si or SiGe devices.

As in Si, the band structure of bulk $Si_{1-x}Ge_x$ has a conduction band at the Δ point when the Ge composition, x, is less than 0.85. The material has a Ge-like character with a conduction band L minima when $x>0.85$ [119]. The energy levels of the conduction band minima and valence band maxima in $Si_{1-x}Ge_x$ alloys vary with composition, which together with strain effects, affect the conduction/valence band offsets of the heterostructure. People *et al.* [14] reported the first theoretical study of the effects of Ge composition on the Si/SiGe band structure: showing that strain causes the conduction and valence bands energy levels to shift and split. Strain in the Si/Ge heterostructure was studied theoretically by Van de Walle and Martin based on self-consistent calculations using local density functional theory and *ab initio* pseudo-potentials [120], considering also the chemical band misalignment and strain effects. Further work on this heterostructure was done by Rieger and Vogl using an empirical pseudopotential method [121]. They concluded that the strain in pseudomorphically strained layers has two main effects on the band structure: the hydrostatic strain shifts the energetic position of a band, and the uniaxial strain component splits degenerate bands [122]. Generally, band offsets enable carrier confinement whereas band splitting affects the effective mass and the density of states, reducing the amount of inter-valley scattering and as a result enhancing carrier transport in the strained layer. To understand the performance enhancements obtainable from a heterostructure it is essential to study the band misalignment and strain effects within the heterostructure.

3.1 Simulation Techniques

Different simulation techniques have been proposed for semiconductor device simulations. A hierarchy of simulation approaches is shown in Fig. 3.1 [123]. These techniques, ranging from complicated numerical intensive quantum approach to the relative simplicity of compact device modelling, provide different levels of description of the semiconductor device and enable wide choices of models for specific

requirements. At the top of the hierarchy are those techniques based on the quantum descriptions of carrier transport. However, the quantum approach is still far from maturity and requires further study. The compact modelling approach at the bottom of Fig. 3.1 uses analytical phenomenological models to describe the device behaviour and is normally used in building blocks for circuit simulations. Therefore, these two models are not of interest in this project and will not be discussed further in this thesis.

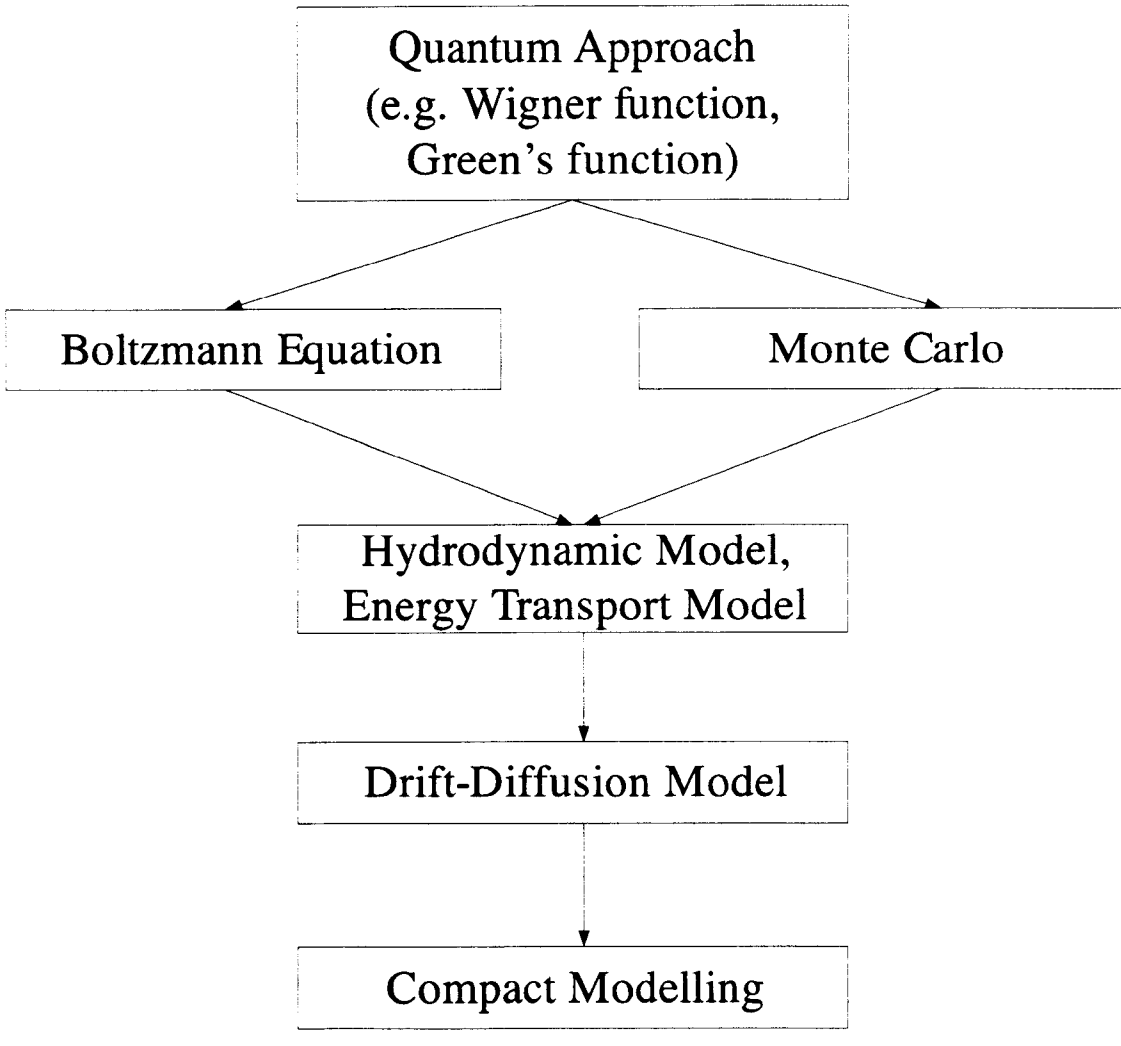


Figure 3.1: Hierarchy of semiconductor simulation models [123]

The backbone of practical numerical device simulations is Poisson's equation, which is discussed in detail in Chapter 4. Poisson's equation relates the electrostatic potential and the electric field to the carrier concentration. To obtain information

on carrier transport it is necessary to relate the carrier velocity to the electric field and the carrier concentration. The semi-classical Boltzmann transport equation, given in equation (3.1), can be used to describe the carrier transport subject the following assumptions [124]:

- The scattering probability is independent of external forces.
- The duration of a collision is much shorter than the average time of motion of a particle.
- External forces are almost constant over a length comparable to the physical dimensions of the wave packet describing the motion of a carrier.
- Band theory and the effective mass theorem apply to the semiconductor under consideration.

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla_r f + \frac{\vec{F}}{\hbar} \cdot \nabla_k f = \left(\frac{\partial f}{\partial t}\right)_c \tag{3.1}$$

Equation (3.1) describes the carrier distribution function f in the seven dimensional phase space which consist of the spatial coordinates $\vec{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, the momentum coordinates $\vec{k} = \begin{bmatrix} k_x \\ k_y \\ k_z \end{bmatrix}$ and time t. $\vec{v}$ is the carrier electron average velocity; $\vec{F}$ is the external force and may be expressed in terms of electric field; $\left(\frac{\partial f}{\partial t}\right)_c$ is the variation of the distribution function due to scattering.

In general it is difficult to solve the Boltzmann equation directly due to the complexity of solving this integro-differential equation in a multi-dimensional space. An equivalent approach is the Monte Carlo simulation approach (shown in Fig. 3.1) which statistically solves the Boltzmann equation without making assumptions of the distribution function. Although this approach can be expensive in terms of computer time, it enables a physical description of the carrier transport in semiconductor devices and has been widely used to determine the real nature of carrier transport.

The Hydrodynamic and Drift-Diffusion approaches, which are discussed in detail in Chapter 4 and Chapter 5, are based on different levels of approximation of the Boltzmann transport equation. The Drift-Diffusion method represents an approximation of the lowest-order transport system obtained from the first momemtum of the Boltzmann transport equation and is suitable for simulating devices close to equilibrium. This means it cannot be used to assess the performance of small devices where non-equilibrium transport maybe important. However, this method is very efficient and provides useful information of device operation, especially the subthreshold behaviour. The details of this method are outlined in Chapter 4.

In additional to the features of the DDM, the Hydrodynamic model conserves momentum and energy and is based on the approximations from the second-order momentum of the Boltzmann equation and accounts for some non-equilibrium transport. However the model which is outlined in detail in Chapter 5 may overestimate the velocity overshoot effect in scaled devices [125]. To precisely describe carrier transport in small devices, Monte Carlo simulation is more appropriate and is discussed in Chapter 6.

Apart from an understanding of the physical models the parameters associated with the Si/SiGe heterostructure are essential for meaningful device simulation in order to properly describe the properties of different materials and evaluate the potential advantages of these new materials. The following sections in this chapter review the most important parameters of the Si/SiGe heterostructure necessary for reliable drift-diffusion and hydrodynamic device simulation. Comparisons between the calculations presented here with published data are given and fitted analytical expressions for those parameters essential for device simulations are provided. The parameters discussed in this chapter include the band gap of strained and relaxed SiGe; the conduction and valence band offsets; the band shifts and splitting due to strain; effective masses; densities of states and permittivity of strained Si and SiGe in the heterostructure.

3.2 Valence Band Offsets in the Si/Ge Heterostructure

One of the most fundamental properties of a semiconductor interface is the band misalignment, *i.e.*, the band offsets between the band edges of the two materials, which determine the effectiveness of carrier confinement. A large conduction band offset is beneficial for electron transport in n-type applications, whereas a large valence band offset is suitable for holes in p-type applications. To date there is less reported work on the Si/SiGe conduction band offset compared to the wide reporting of the theoretical and experimental studies of valence band offsets [120 - 122,126 132]. The band offsets depend on the crystallographic orientation [120] and only the commonly studied (100) orientation Si/SiGe heterostructure is discussed in this chapter.

The first theoretical study of the valence band offset of Si/SiGe heterostructure was based on the calculation of the offset between pure Si and Ge, for which Van de Walle and Martin proposed two approaches: one is the self-consistent interface calculation [120] and the other is the "model solid" approach [126]. The theoretical studies predicted valence band offsets of 0.84 eV [120], 0.93 eV [126] and 0.74 eV [127] for pure Ge on Si, and 0.31 eV [120], 0.36 eV [126] and 0.21 eV [127] for pure Si on Ge. These may be compared to the experimental data of 0.74±0.13 eV [128], 0.83±0.11 eV [129] and 0.78 eV [130] for pure Ge on pure Si, and 0.17±0.13 eV [128] and 0.22±0.13 eV [129] for pure Si on pure Ge.

The linear interpolation of the valence band offset in a Si/Ge heterostructure with respect to Ge composition has been adopted to calculate the valence band offsets of the $Si/Si_{1-x}Ge_x$ or $Si_{1-x}Ge_x/Si$ systems [14, 120, 127, 131]. Moreover, when the hydrostatic and uniaxial strain components are considered separately, published results show that the average valence band discontinuity is nearly independent of the strain conditions and only varies with a weak linear relationship on the lattice constant parallel to the interface [120,127,130]. Neglecting the weak linear variation of the average valence band due to strain, Yu *et al.* [129] reported a 0.49±0.13 eV discontinuity of the average valence band edges for pure Ge on Si from their x-ray

photoelectron spectroscopy measurements. Using self-consistent interface calculations [120] and the "model solid" [126] approach, Van de Walle and Martin predicted 0.54 eV and 0.63 eV respectively for the average valence band offset of pure Ge on Si, and 0.53 eV and 0.60 eV respectively for pure Si on Ge. Following the calculations by Colombo *et al.* [127], whose results agreed with the core-level photoemission data from Schwartz *et al.* [128], Rieger and Vogl [121] proposed an interpolation formula for the average valence band offset (due to the effects of hydrostatic strain), $\triangle E_{v,av}$ at the interface between a strained $Si_{1-x}Ge_x$ layer on a relaxed $Si_{1-y}Ge_y$substrate:

$$\triangle E_{v,av} = (0.47 - 0.06y)(x - y) \tag{3.2}$$

A realistic measure of the valence band offset, $\triangle E_v$, may be found by applying the effect of the uniaxial strain components on $\triangle E_{v,av}$. The lowest conduction band edge offset, $\triangle E_c$, is then calculated by subtracting $\triangle E_v$ from the difference between the bandgaps of the active strained layer and the relaxed substrate layer. Note that the symbol, $\triangle E_{c(v)}$, used here denotes the energy level difference between the active and substrate layers: $\triangle E_{c(v)}$=$E_{c(v)}$(active)-$E_{c(v)}$(substrate).

3.3 The Bandgap of the Unstrained $Si_{1-y}Ge_y$ and Strained $Si_{1-x}Ge_x$ Alloy

The $Si_{1-y}Ge_y$ alloy has been found to have an indirect band gap and the conduction band changes from Si-like with six-fold ellipsoid Δ-valley minima (shifted by 15% from X-valley) to Ge-like with L-valley minima when the Ge content y>0.85, see Fig. 3.2 [119, 133, 134]. Braunstein *et al.* [119] first studied the dependence of the bandgap on the Ge content and temperature from optical absorption measurements. The dependence of the $Si_{1-y}Ge_y$ bandgap on Ge content was confirmed by Krishnamurithy *et al.* [134] using the virtual crystal approximation (VCA) and the molecular coherent-potential approximation (MCPA). Good agreement was obtained with experimental data. Weber *et al.* [133] measured the $Si_{1-y}Ge_y$ bandgap using photoluminescence (PL) at 4.2K from which the following fitted quadratic

expression as a function of Ge content, y, was obtained [133] for $y<0.85$:

$$E_g(y) = 1.155 - 0.43y + 0.206y^2 \tag{3.3}$$

Although all efforts predict similar trends of the bandgap with variation of the Ge content, from Fig. 3.2 deviations of up to 40 meV between data reported by Braunstein *et al.* [119] and Weber *et al.* [133] are observed . However, equation (3.3) tends towards the accepted value for Si at low temperature and is adopted to estimate the bandgap of unstrained $Si_{1-y}Ge_y$ as a function of y at 4.2K.

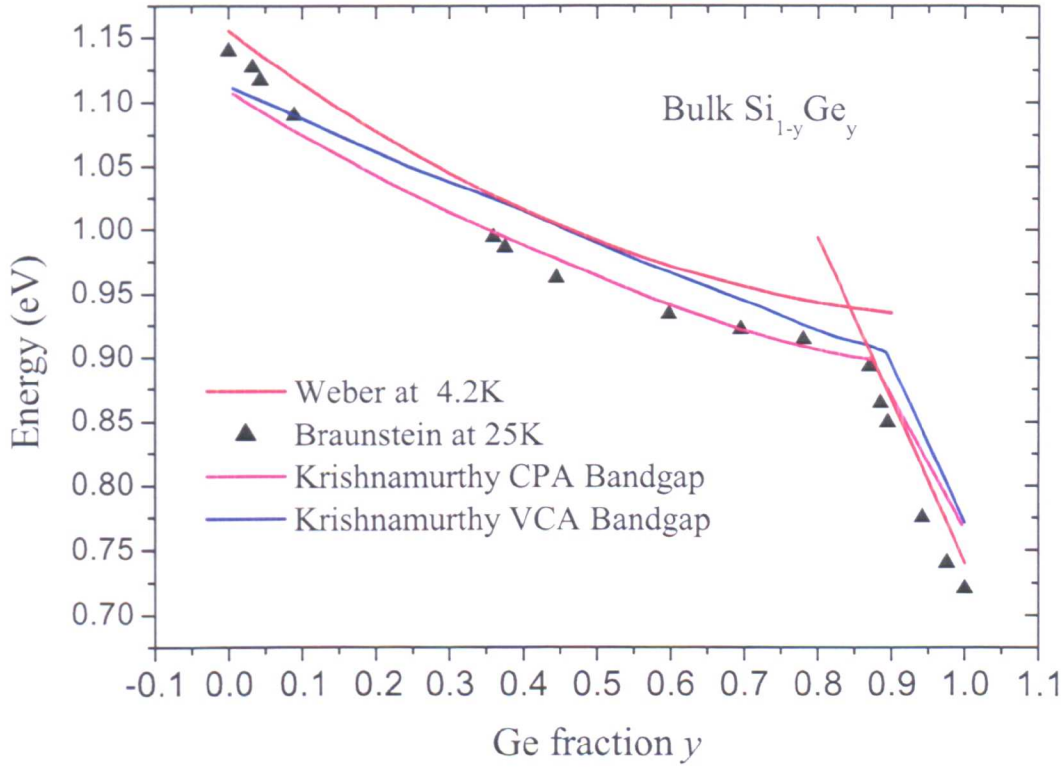


Figure 3.2: Bandgap of bulk $Si_{1-y}Ge_y$ alloy

To estimate the $Si_{1-y}Ge_y$ bandgap at higher temperatures, one may use the well-known relationship for the temperature (T) dependence of the bandgap of Si [135] given below:

$$E_g(T) = E_g(0K) - \frac{4.73 \times 10^{-4} T^2}{T + 636} \tag{3.4}$$

Fig. 3.3 compares the bandgap for a given Ge content as a function of temperature between Braunstein *et al.* [119] (symbols) and the calculations using equations (3.3) and (3.4) (Note that the energy gap at *0K* is approximated by the value at 4.2K). The agreement indicates that the temperature dependence of the bandgap of unstrained $Si_{1-y}Ge_y$ may be estimated using the same temperature dependence as for Si.

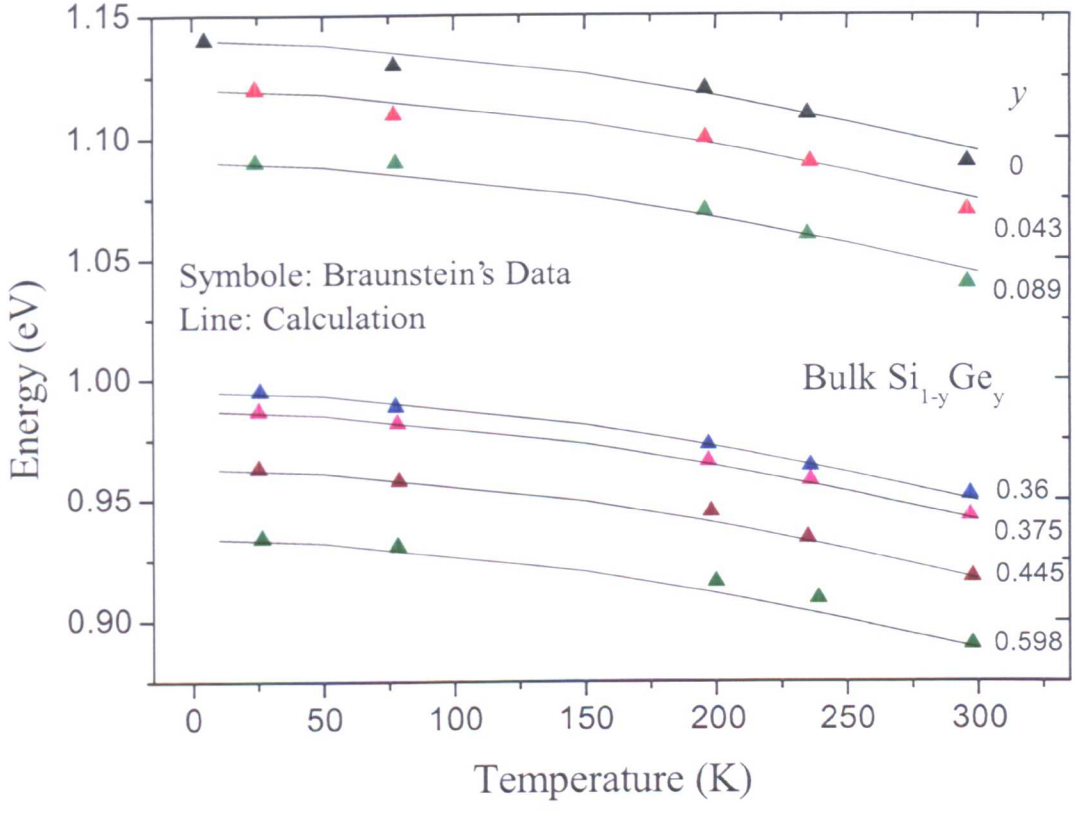


Figure 3.3: Temperature dependence of the bandgap for $Si_{1-y}Ge_y$ bulk alloys

There have been many investigations of the bandgap of pseudomorphically grown strained $Si_{1-x}Ge_x$ on a relaxed Si substrate. Fig. 3.4 compares the experimental data and some analytical results for the bandgap of low temperature strained $Si_{1-x}Ge_x$ as a function of Ge content x [136–141]. For Ge contents of up to 30%, there is good agreement between experimental data and the analytical work. At Ge contents up to 60%, the expressions obtained by Robbins *et al.* (upper triangle symbols in Fig. 3.4) from photoluminescence measurements [136] and Bean *et al.* (lower triangle symbols

in Fig. 3.4) [138] deviate by a maximum of 60 meV from Lang *et al.*'s absorption measurement data [140]. However, in addition to the fundamental bandgap, Lang's data includes the average energy of phonons involved in the absorption process. Therefore, by adding the exciton binding energy to the experimentally determined energy of the free exciton line (upper triangle symbols in Fig. 3.4), the bandgap between the strained $Si_{1-x}Ge_x$ valence band and the lowest conduction band edge at 4.2 K (for $x<0.4$) may be estimated [140] (square symbols in Fig. 3.4) by equation (3.5) shown below:

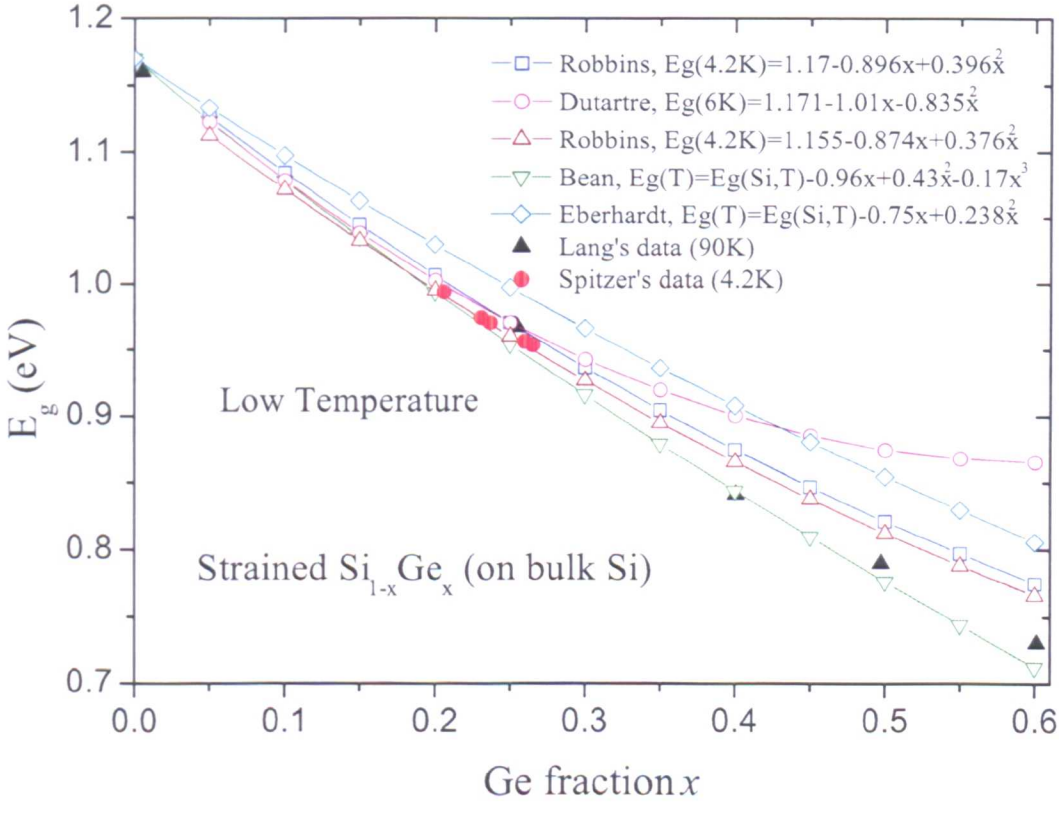


Figure 3.4: Bandgap of strained $Si_{1-x}Ge_x$ on Si (001) substrate

$$E_g(x) = 1.17 - 0.896x + 0.396x^2 \tag{3.5}$$

The bandgap temperature dependence relationship for Si may be used to estimate the bandgap of strained $Si_{1-x}Ge_x$ [139] at higher temperatures (equation (3.4)).

3.4 Strain Effects on the Band Structure

For Ge contents less than 85%, the $Si_{1-x}Ge_x$ alloy exhibits Si-like band structure, having six-fold degenerate conduction band minima and two-fold degenerate valence band maxima. The lattice constant mismatch between the pseudomorphic $Si_{1-x}Ge_x$ active layer and the relaxed $Si_{1-y}Ge_y$ substrate material determines the strain type: which is compressive if $x>y$ or tensile if $x>y$. The strain in the pseudomorphic active layer includes a hydrostatic component which shifts the average band energy level and an uniaxial component which splits the degenerate bands. Here, symbol $\triangle E$ is used to denote the energy level shift induced by the strain: $\triangle E - E$(strained)-E(relaxed).

The relaxed $Si_{1-y}Ge_y$ lattice constant as a function of Ge content, y, is given by [122]:

$$a_0(y) = a_0(Si) + 0.200326y(1-y) + [a_0(Ge) - a_0(Si)]y^2 \tag{3.6}$$

where $a_0(Ge)$ and $a_0(Si)$ are the lattice constants of relaxed Ge and Si, respectively. Using $a_\perp$ and $a_\parallel$ to represent the parallel (in-plane) and perpendicular (to the interface) lattice constants in the strained $Si_{1-x}Ge_x$ layer, then we have [121]:

$$\begin{cases} a_\parallel = a_0(y) \\ a_\perp = a_0(x)\left[1 - 2\frac{c_{12}(x)}{c_{11}(x)}\frac{a_\parallel - a_0(x)}{a_0(x)}\right] \end{cases} \tag{3.7}$$

where $c_{11}(x)$ and $c_{12}(x)$ are the elastic constants for the $Si_{1-x}Ge_x$ alloy found from the linear interpolation between those of Si and Ge. The lateral, $\varepsilon_\parallel$, and perpendicular, $\varepsilon_\perp$, strain tensors in strained $Si_{1-x}Ge_x$ are then defined as [121]:

$$\begin{cases} \varepsilon_\parallel = \frac{a_\parallel}{a_0(x)} - 1 \\ \varepsilon_\perp = \frac{a_\perp}{a_0(x)} - 1 \end{cases} \tag{3.8}$$

The hydrodynamic strain, $\triangle V/V$, is related to the hydrostatic strain shifts of the conduction and valence bands by following expressions [122]:

$$\triangle E_{c,av} = a_c\frac{\triangle V}{V} = a_c(2\varepsilon_\parallel + \varepsilon_\perp) \tag{3.9}$$

$$\triangle E_{v,av} = a_v \frac{\triangle V}{V} = a_v(2\varepsilon_{\parallel} + \varepsilon_{\perp}) \tag{3.10}$$

where a_c and a_v are the hydrodynamic deformation potentials for the conduction band and valence band respectively.

Under the action of uniaxial strain along the [001] direction, the bands along [100] and [010] directions split off from the one along the [001] direction. The uniaxial induced energy shifts of the conduction bands are given by [122]:

$$\triangle E_c^{001} = \frac{2}{3}\Xi_u^{\triangle}(\varepsilon_{\perp} - \varepsilon_{\parallel}) \tag{3.11}$$

$$\triangle E_v^{100,010} = -\frac{1}{3}\Xi_u^{\triangle}(\varepsilon_{\perp} - \varepsilon_{\parallel}) \tag{3.12}$$

where Ξ_u is the uniaxial strain deformation potential for the conduction band. The energy splitting of the valence bands is given by [122]:

$$\triangle E_{v,1} = -\frac{1}{6}\triangle_0 + \frac{1}{4}\delta E + \frac{1}{2}\left[\triangle_0^2 + \triangle_0\delta E + \frac{9}{4}\delta E^2\right] \tag{3.13}$$

$$\triangle E_{v,2} = -\frac{1}{3}\triangle_0 - \frac{1}{2}\delta E \tag{3.14}$$

$$\triangle E_{v,3} = -\frac{1}{6}\triangle_0 + \frac{1}{4}\delta E - \frac{1}{2}\left[\triangle_0^2 + \triangle_0\delta E + \frac{9}{4}\delta E^2\right] \tag{3.15}$$

where $\triangle_0$ is the spin-orbit splitting; for strain along the [001] direction, $\delta E = 2b$(and b is the uniaxial deformation potential for tetragonal strain.

The parameters are summarized in table 3.1 [120 122, 142]. Note that these parameters are for pure Si and Ge. Linear interpolation may be used to calculate the parameters for $Si_{1-x}Ge_x$ alloys according to the Ge content.

Table 3.1: The parameters for the calculation of the strain effects on the band structure

Parameters at 300K	Si	Ge
Bandgap $E_g(eV)$	1.12	0.66
Lattice constant (Å)	5.43	5.65
$c_{11}(Mbar)$	1.675	1.315
$c_{12}(Mbar)$	0.65	0.494
$a_c(eV)$	4.18	2.55
$\Xi_u(eV)$	9.16	9.42
$a_v(eV)$	2.46	1.24
$\triangle_0(eV)$	0.044	0.296
l (eV)	-0.9	-2.4
m (eV)	3.6	4.2
n (eV)	-5.889	-7.621

Fig. 3.5 illustrates the tensile strain induced energy level shifts and the effect of splitting on the Si band structure. The hydrostatic tensile strain raises the average energy levels of the conduction band and valence band and the uniaxial component splits degenerate bands. This causes a lowering of the twofold degenerate [001] ellipsoids and raises the fourfold degenerate [100][010] ellipsoids from the sixfold degenerate system.

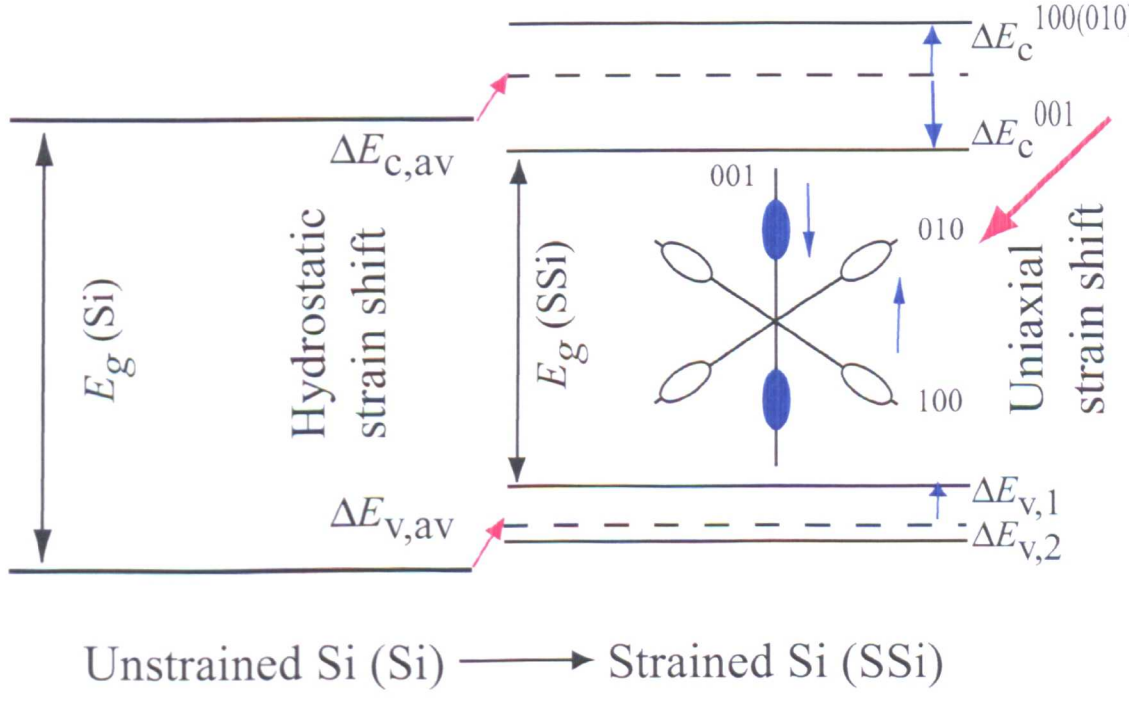


Figure 3.5: The strain effects to the band structure of silicon

The strain also shifts and splits the degenerate valence bands at the Γ points (see Fig. 3.5). Fig. 3.6 shows the calculated splitting of the valence band into levels *v1*, *v2* and *v3*. Compressive strain raises band *v2* (which has a larger effective mass) whereas tensile strain will raise band *v1*. Note that the split valence band *v2* is a pure |3/2,3/2> state, whilst bands *v1* and *v3* are the mixtures of states |3/2,1/2> and |1/2,1/2>. The splitting of the conduction and valence bands affects the densities of states and the effective masses; these will be discussed later in this chapter.

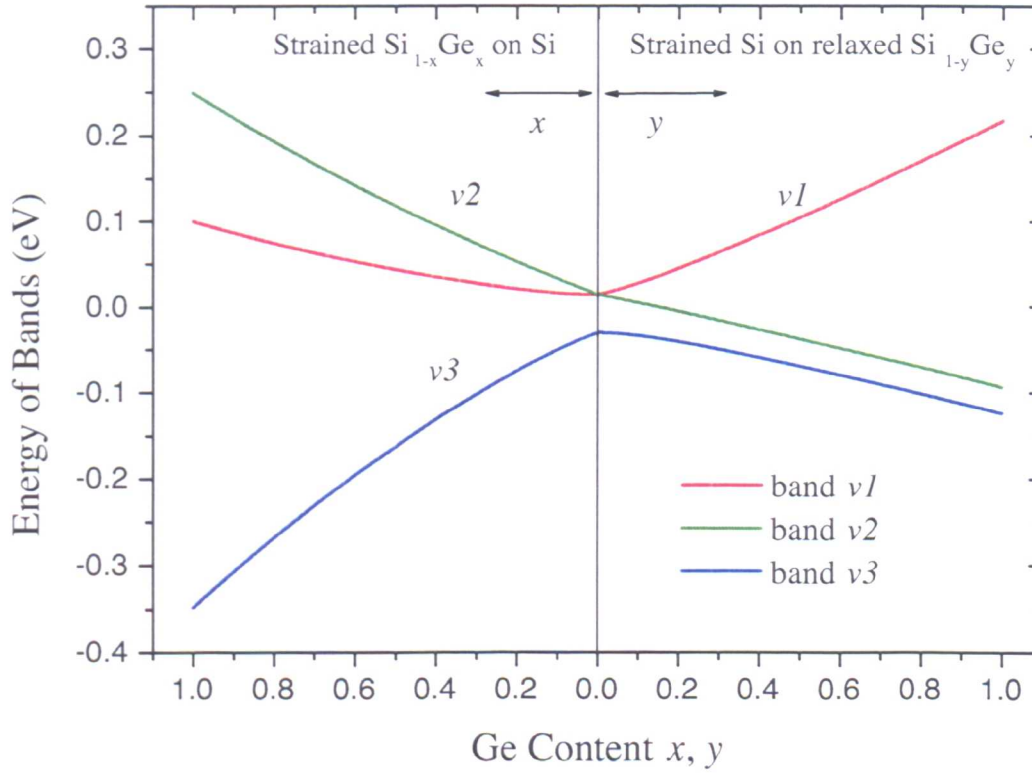


Figure 3.6: The valence band splitting (taking the average valence band as the reference level)

The bandgap of strained Si may then be calculated by subtracting the total energy level shifts of the lowest conduction band level and the highest valence band level from the unstrained Si bandgap (see Fig. 3.5). A linear relationship for the strained Si bandgap as a function of Ge content has been deduced from [14] and is plotted together with experimental data from [121] in Fig. 3.7. However, further

experimental studies are necessary to determine the strained Si bandgap with large Ge contents within the SiGe substrate. The linear fit (solid line in Fig. 3.7) to the calculations (square symbols in Fig. 3.7, with a maximum error of 3%) for the bandgap of strained Si as a function of the Ge content, y, of the relaxed $Si_{1-y}Ge_y$ substrate at 300 K is given by:

$$E_g(y) = 1.11 - 0.6y \tag{3.16}$$

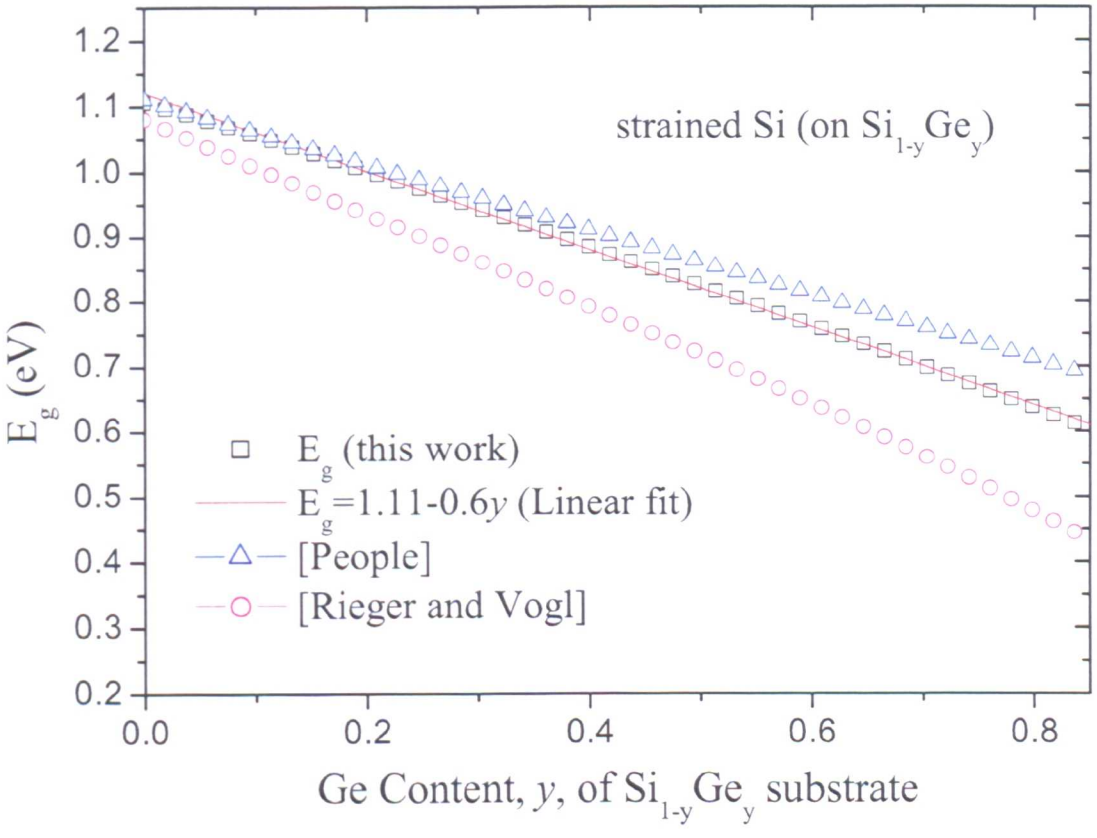


Figure 3.7: Calculated bandgap of strained Si on relaxed $Si_{1-y}Ge_y$ as a function of Ge content y, compared with the calculations taken from People [14] and Rieger *et al* [121]

3.5 Band Offsets in the $Si/Si_{1-x}Ge_x$ Heterostructure

3.5.1 The strained $Si_{1-x}Ge_x$ on relaxed Si heterostructure

The type I strained $Si_{1-x}Ge_x$ on unstrained Si heterostructure has a large valence band offset which enables enhanced hole transport, see Fig. 2.1. It has been widely used in SiGe HBT and p-type MOSFET applications. The compressive strain in the pseudomorphic $Si_{1-x}Ge_x$ layer causes the lowering of the fourfold degenerate ellipsoids and the raising of the twofold degenerate ellipsoids (*cf.* the effect of tensile strain) in the conduction band and also splits the degenerate valence bands.

Using equation (3.2), the average offset between the valence bands of strained $Si_{1-x}Ge_x$ and Si is calculated by setting y—0, giving $\triangle E_{v,av} = 0.47x$. This allows us to use the Si valence band edge as the reference energy level for the following band offset calculations of the $Si_{1-x}Ge_x/Si$ heterostructure. The strain effects on the $Si_{1-x}Ge_x$ band structure can be calculated by using the theory described earlier in this paper, which gives both the hydrodynamic strain and the uniaxial band splitting components. Taking the valence band splitting into account, the valence band offset $\triangle E_v$, which is the offset between the highest valence band edges of the heterostructure, may be estimated by:

$$\triangle E_v = \triangle E_{v,av} + max(\triangle E_{v1}, \triangle E_{v1}, \triangle E_{v1}) = 0.71x \tag{3.17}$$

Note that the linear expression is in good agreement with experimental data [132] for Ge contents less than 30%. For example, for pure Ge on Si, this linear relationship gives a 0.71 eV valence band offset which is close to the reported value of 0.74 eV [127, 128]. The calculated offsets of the highest valence band edge and lowest conduction band edge between the strained $Si_{1-x}Ge_x$ and Si are given in Fig. 3.8, compared with theoretical results [120] and experimental data [132].

The conduction band offset, $\triangle E_c$, is found by subtracting the Si bandgap from the sum of the calculated valence band offset $\triangle E_v$ and the $Si_{1-x}Ge_x$ bandgap given by equation (3.5):

$$\triangle E_c = E_g(SSiGe) + \triangle E_v - E_g(SSi) \tag{3.18}$$

However, the conduction band offset, which varies nonlinearly with Ge content, is rather less than the valence band offset and is generally not of interest in this type of heterostructure.

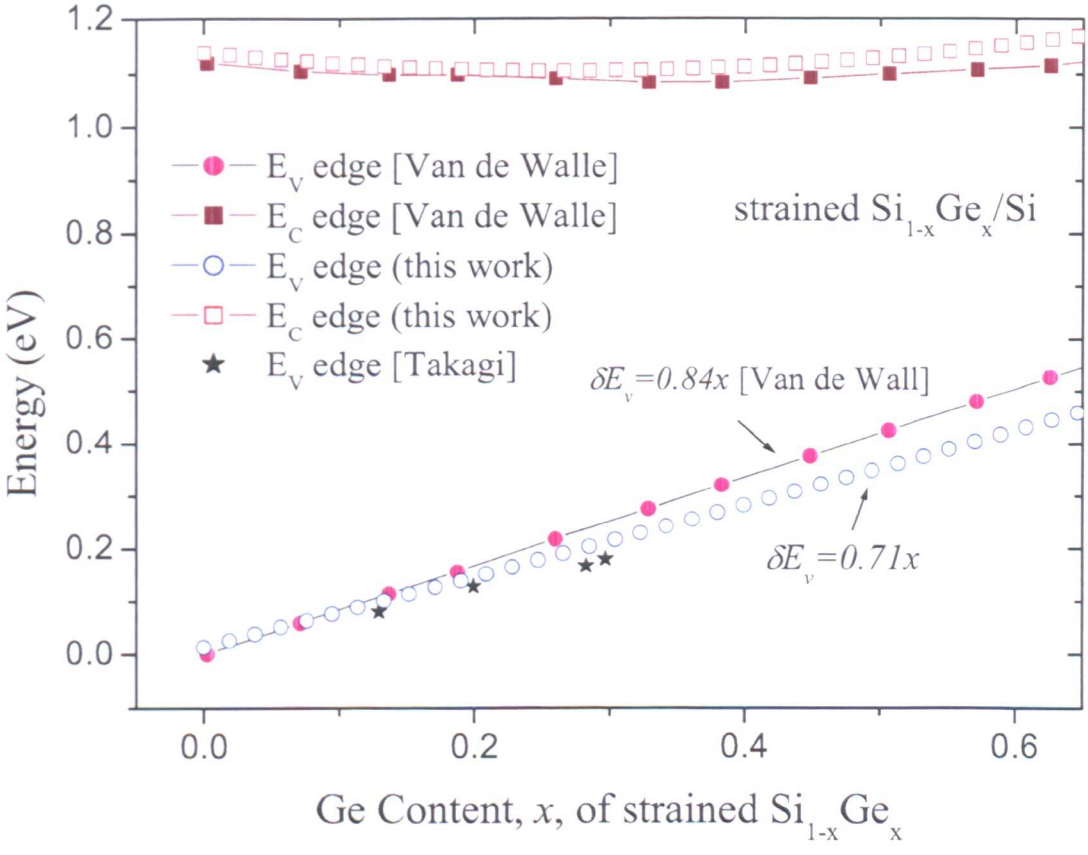


Figure 3.8: Calculated valence and conduction band offsets between strained $Si_{1-x}Ge_x$ and unstrained Si (assuming the valence band edge of Si as the reference zero of the energy level), compared with the data taken from Walle *et al.* [120] and Takagi *et al.* [132]

3.5.2 The strained Si on $Si_{1-y}Ge_y$ heterostructure

The type II strained Si on $Si_{1-y}Ge_y$ heterostructure has considerable band offsets in both the conduction and valence bands. The strain induces band splitting of both degenerate conduction and valence bands, lowers the effective mass and reduces the

inter-valley scattering and therefore enhances both hole and electron transport.

The calculation of the heterostructure band offsets are based on the average offset of the valence band, $\triangle E_{v,av} = (0.06y - 0.47)y$, obtained from rearrangement of equation (3.2). The valence band offset, $\triangle E_v$, is determined by considering the highest valence band splitting of $\triangle E_{v,av}$. The conduction band offset, $\triangle E_c$, is then calculated by subtracting the bandgap of relaxed $Si_{1-y}Ge_y$, given by equation (3.4), from the sum of the strained Si bandgap, given by equation (3.16), and the valence band offset of the heterostructure, $\triangle E_v$.

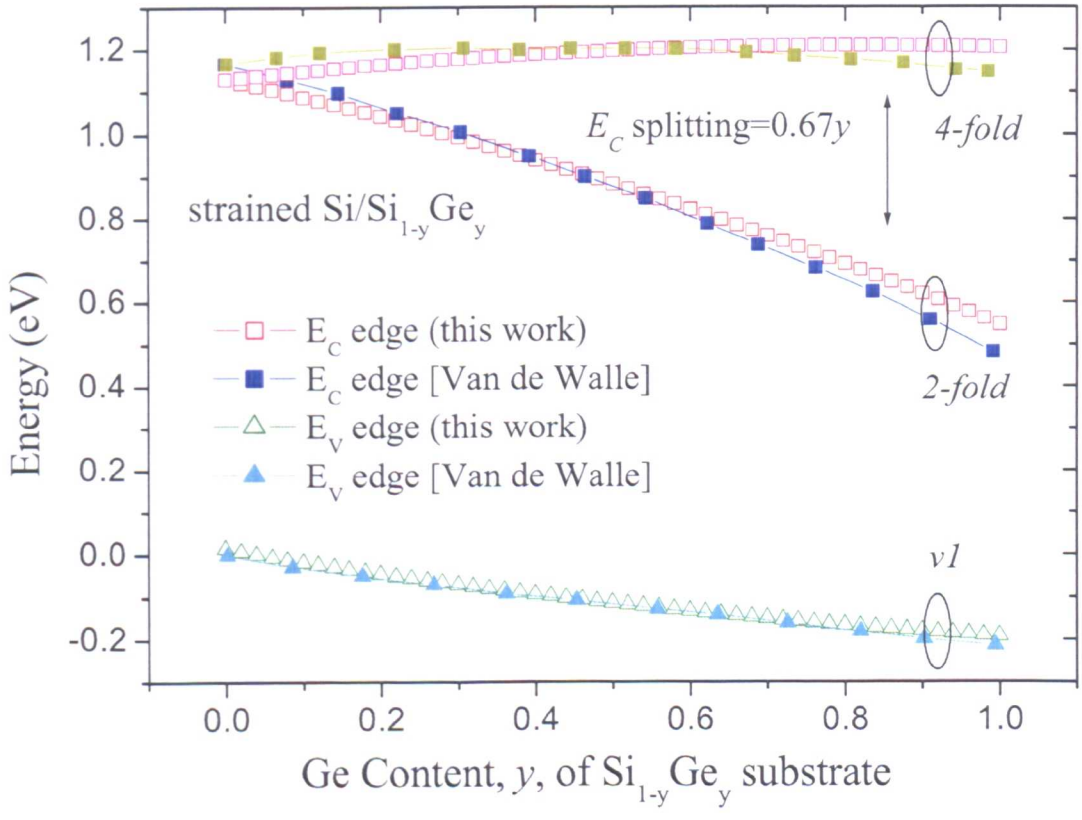


Figure 3.9: Calculated conduction band splitting and the valence band edge of strained Si on relaxed $Si_{1-y}Ge_y$ as a function of Ge composition y (assuming the valence band edge of Si as the reference zero of the energy level), compared with the work by Walle *et al.* [120] and Rieger *et al.* [121]

Fig. 3.9 shows the calculated conduction band splitting and the highest valence band edge as a function of Ge content, together with the theoretical calculations from [120, 121] for comparison. The valence band offset, $\triangle E_v$, may be expressed by the following quadratic expression as a function of Ge content, y,

$$\triangle E_v = -0.238y + 0.03y^2 \tag{3.19}$$

where the negative sign indicates that the valence band of strained Si is lower than that of relaxed $Si_{1-x}Ge_x$. At $y{-}1$, valence band offset of 0.208 eV is obtained for pure Si on Ge, which is close to the value of 0.22 eV given by [127,129]. The conduction band splitting of strained Si exhibits a linear relationship, $\triangle E_{c,splitting} = 0.67y$ [14]. It is found that the following cubic fit for the conduction band offset as a function of the substrate Ge content, y, can be used.

$$\triangle E_c = E_g(SSi) + \triangle E_v - E_g(SiGe) = -0.35y - 0.35y^2 + 0.12y^3 \tag{3.20}$$

In equation (3.20), the negative offset denotes that the conduction band edge of strained Si is lower than that of relaxed $Si_{1-y}Ge_y$, providing electron confinement in the strained Si layer at the strained $Si/Si_{1-y}Ge_y$ interface.

3.6 Other Parameters

3.6.1 Effective Mass

The effective electron mass in the strained $Si_{1-x}Ge_x$ or relaxed $Si_{1-y}Ge_y$ alloy has been calculated by Rieger and Vogl [121], where they derive the effective longitudinal and transverse electron masses in strained $Si_{1-x}Ge_x$ on relaxed $Si_{1-y}Ge_y$ as functions of x and y. Under tensile strain the conduction band minimum of strained $Si_{1-x}Ge_x$ is located in the lowest doubly degenerate band and in the lowest fourfold degenerate band for compressive strain. Therefore, at Ge contents where the large degree of conduction band splitting reduces the occupation in the raised subvalleys, the effective electron mass of strained Si layer or strained SiGe layer may be assumed to be that associated with the lowest twofold or fourfold band minima. The analytical expressions for the longitudinal and transverse electron effective masses can be obtained from [121] and will not be given in this chapter.

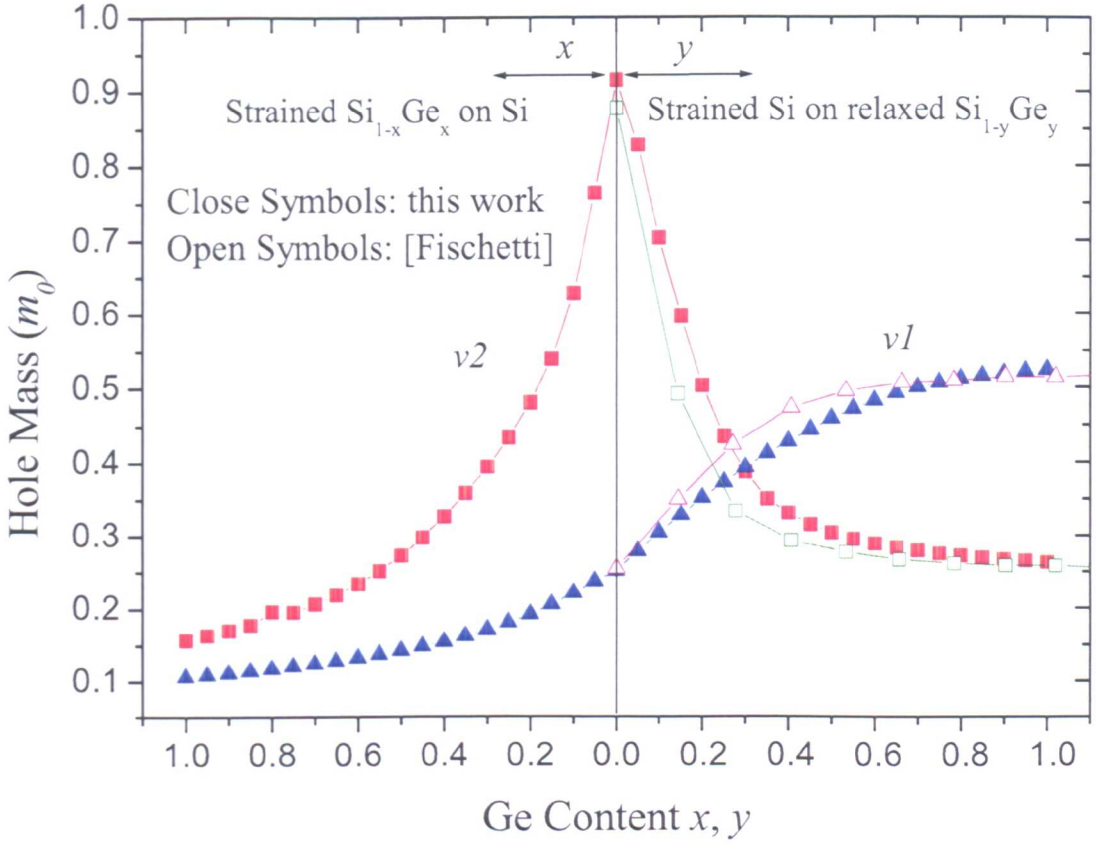


Figure 3.10: Calculated hole effective mass in band *v1* and *v2* from full-band $k \bullet p$ calculations, compared with data from Fischetti *et al.* [142]

In this work, the effective hole masses are calculated directly from six band $\boldsymbol{k \bullet p}$ theory using in-house ensemble full-band Monte Carlo simulator [1]. The calculated hole masses in bands *v1* and *v2* of strained Si (on relaxed $Si_{1-y}Ge_y$) and strained $Si_{1-x}Ge_x$ (on Si), have been given in Fig. 3.10. For $Si_{1-x}Ge_x$ under compressive strain, the mass in *v2* is larger than that in *v1* and from Fig. 3.6 it is evident that band *v2* is raised during the splitting. Therefore, in strained $Si_{1-x}Ge_x$ band *v2* plays an important role in hole transport. Under tensile strain, the mass in *v2* is larger than that in band *v1* for Ge contents less than 20%, whereas the masses in band *v1* are greater than those in *v2* for greater Ge contents. However, it is clear from Fig. 3.6 that band *v1* is at the top of the valence band edge and therefore both bands will contribute to the hole transport at low Ge contents. With increasing Ge

[1] Courtesy of Dr. J. R. Watling, Device Modelling Group, University of Glasgow

content, band *v1* dominates transport due to the higher mass and the increasing band splitting reduces the occupation of band *v2*. The calculated masses for the tensile strained Si have been compared with the results from [142] in Fig. 3.10. The slight disagreements are due to the different deformation potentials used in the calculations.

The calculated masses for the two strained materials have been fitted to the analytical expressions below for strained $Si_{1-x}Ge_x$ (equation (3.21)) and strained Si on relaxed $Si_{1-y}Ge_y$ (equation (3.22)), which show good agreement for Ge contents up to 80%.

$$\begin{bmatrix} m(SSiGe,v2) \\ m(SSiGe,v1) \end{bmatrix} /m_0 = \begin{bmatrix} -2.8369 \\ -0.1432 \end{bmatrix} x^3 + \begin{bmatrix} 4.6844 \\ 0.3618 \end{bmatrix} x^2 + \begin{bmatrix} -2.87 \\ -0.3699 \end{bmatrix} x + \begin{bmatrix} 0.8956 \\ 0.2534 \end{bmatrix} \tag{3.21}$$

$$\begin{bmatrix} m(SSi,v2) \\ m(SSi,v1) \end{bmatrix} /m_0 = \begin{bmatrix} -2.2515 \\ 0.0298 \end{bmatrix} y^3 + \begin{bmatrix} 4.4158 \\ -0.3308 \end{bmatrix} y^2 + \begin{bmatrix} -2.9248 \\ 0.5713 \end{bmatrix} y + \begin{bmatrix} 0.9375 \\ 0.2511 \end{bmatrix} \tag{3.22}$$

3.6.2 Density of States

For strained $Si_{1-x}Ge_x$ on relaxed $Si_{1-y}Ge_y$, the conduction band splitting, $\triangle E_{c,splitting}$ between the lowered a-fold degenerate bands and the raised b-fold degenerate bands may be determined from equations (3.11) and (3.12); where for tensile strain, a=2 and b=4 and in the case of compressive strain, a=4 and b=2. The free carrier density can then be calculated by

$$n = 2\int_{E_c}^{\infty} (a + b\exp(-\frac{\triangle E_{c,splitting}}{k_B T}))f_0(E)N(E)dE \tag{3.23}$$

where $f_0(E)$ is normally taken to be Maxwell-Boltzmann approximation of the Fermi-Dirac statistic. *N(E)dE* is the density of states per unit volume in the energy range E to $E+dE$. The multiplier 2 outside of the integral accounts for spin and the factors inside the integral represent the degeneracies of the lowered a-fold degenerate

band and the raised b-fold degenerate band respectively. This equation may then be cast in the familiar form $n = N_c \exp\left(-\frac{E_c - E_F}{k_B T}\right)$. The effective density of states in conduction band, N_c is therefore defined by

$$N_c = 2(a + b\exp(-\frac{\triangle E_{c,splitting}}{k_B T})) \times (\frac{m k_B T}{2\pi\hbar^2})^{3/2} \tag{3.24}$$

where m is calculated from [121]. In typical applications (where the Ge content is greater than 13%), the exponential term in (3.24) is around 0.05, which enables us to neglect the contribution from the raised b-fold bands. Therefore, the effective conduction band density of states of strained SiGe is approximated as the effective DOS of the lowest a-fold degenerate valleys, as illustrated in Fig. 3.11. For Ge contents greater than 15%, the conduction band DOS of tensile strained Si is nearly 1/3 that of relaxed Si, while the conduction band DOS of compressive strained SiGe drops to 2/3 of the DOS of relaxed Si. However, there is a slight increase with increasing Ge content which is due to the increase in the effective electron mass.

In the valence band, strain causes the splitting of degenerate valence bands at the Γ point. The strain shifts bands *v1* and *v2* up for compressive strain and band *v1* up and band *v2* down for tensile strain (see Fig. 3.6). This band splitting reduces the occupation in the lowered subvalleys and therefore decreases the density of states. Using equations (3.21) and (3.22) to calculate the hole mass in each band, the density of states in the valence band, N_v can be obtained using an analysis analogous for N_c:

$$N_v = 2(m_U^{3/2} + m_M^{3/2} \exp(-\frac{\triangle E_{v,splitting}}{k_B T})) \times (\frac{k_B T}{2\pi\hbar^2})^{3/2} \tag{3.25}$$

where m_U and m_M are the effective hole masses of the raised and lowered bands respectively and $\triangle E_{v,splitting}$ is the valence band splitting of the degenerate band at the Γ point. Note that equation (3.25) neglects the contribution from band *v3* (spin split-off band) which is relatively far from the top valence band edge (see Fig. 3.6). The density of states in the conduction and valence bands as a function of Ge content is given in Fig. 3.11. The upturn in the strained Si valence band density of

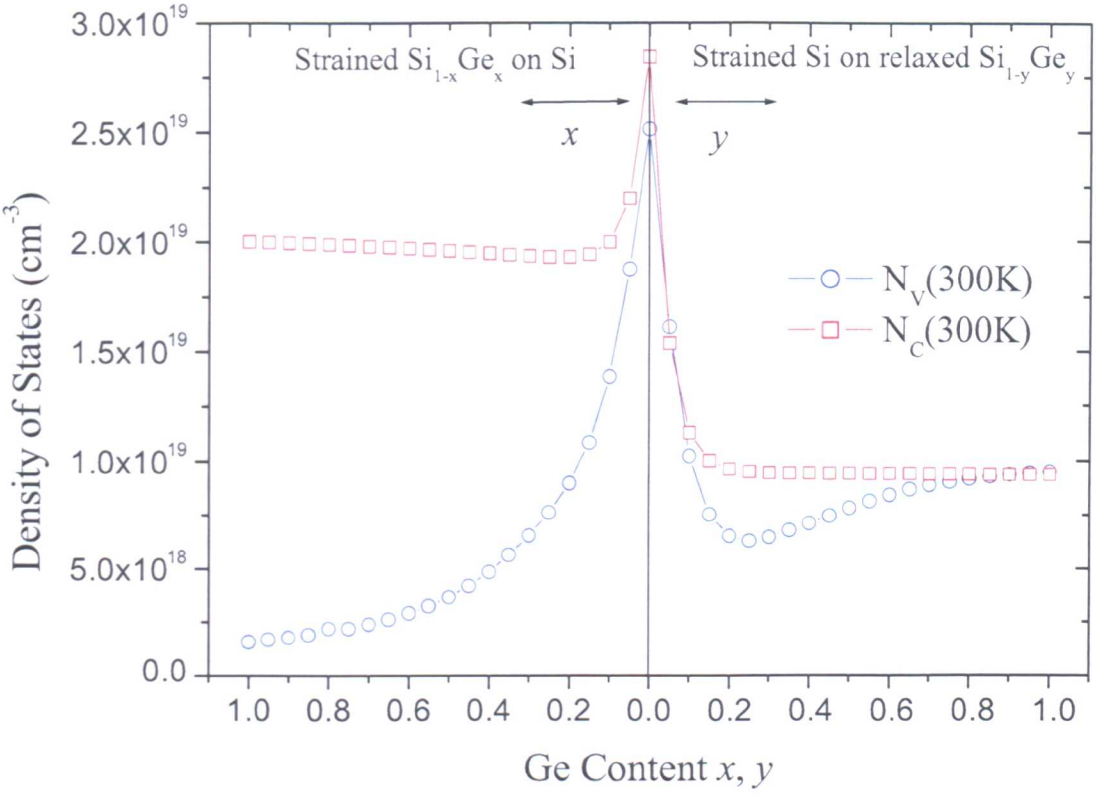


Figure 3.11: Calculated electron and hole density of states of tensile strained Si and compressive strained SiGe at 300K

states is due to the increased hole mass in band *v1* and becomes larger than that in band *v2* when the Ge content is greater than 20%. It can be found from equation (3.25) that N_v reverts to the case of relaxed Si with its larger density of states (see Fig. 3.11) compared to the values given in [135,143]. However, the density of states given by [135,143] uses a low temperature mass ($0.5m_0$) for the heavy hole that is inappropriate at room temperature [29].

3.6.3 Permittivity

An alternative interpolation scheme to estimate the permittivity of the $Si_{1-x}Ge_x$ alloy based on the Clausius-Mossotti relationship [144] that gives:

$$\frac{\epsilon_{SiGe}-1}{\epsilon_{SiGe}+2} = (1-x)\frac{\epsilon_{Si}-1}{\epsilon_{Si}+2} + x\frac{\epsilon_{Ge}-1}{\epsilon_{Ge}+2} \tag{3.26}$$

Based on this relationship, the following formula can be used for simplified com-

putation:

$$\epsilon_{SiGe} = \frac{2(A+B)+1}{1-(A+B)} \tag{3.27}$$

where $A = (1-x)\frac{\epsilon_{Si}-1}{\epsilon_{Si}+2}$ and $B = x\frac{\epsilon_{Ge}-1}{\epsilon_{Ge}+2}$.

3.7 Summary

This chapter summarizes relevant results related to the bandgap and band misalignment calculations of Si/SiGe heterostructure, focusing on their application in practical numerical simulation of related devices. The aim is to provide simple empirical expressions for the majority of essential parameters associated with transport in this type of heterostructure and complemented with guidance for their implementation in corresponding semi-classical device simulations. By comparing the existing theoretical and experimental work with these calculations, analytical expressions have been concluded for the bandgap; the band offset; the effective masses; the densities of states and the permittivity for the strained Si on relaxed $Si_{1-y}Ge_y$ and the strained $Si_{1-x}Ge_x$ on unstrained Si heterostructure.

Chapter 4

Simulations of Strained-Si/SiGe MODFETs for RF and High Linearity Applications

This chapter focuses on strained Si MODFETs for high frequency and high linearity applications. Similar to their III-V counterparts, Si/SiGe MODFETs have channels with high mobility and high sheet carrier density. Moreover, the compatibility of Si/SiGe MODFETs with the existing Si technology makes them promising candidates for system-on-chip applications. However, the complicated layer structure of MODFETs requires proper design to achieve high device performance, as discussed in Chapter 2. Compared to previous work on the optimization of the device structure for RF applications, little work has been done to optimize such devices for communication applications, bearing in mind that power amplification for wide-band communications requires high linearity to minimize the intermodulation distortion.

Comprehensive numerical simulations are used to study the effects of both lateral and vertical device designs on device performance and linearity. The simulations are based on extensive calibrations in respect of a 0.25 μm and a 70 nm n-type buried strained Si channel Si/SiGe MODFETs fabricated by Daimler Chrysler. Apart from the study of different device geometries, the impact of different doping strategies on device performance and linearity behaviour are also investigated.

4.1 Simulation Tools and Models

The simulation work in this chapter uses 2-dimensional (2-D) drift-diffusion model (DDM) device simulators, MEDICI and TAURUS from Synopsys [27], and Greg-Snider 1-D Poisson-Schrödinger solver [28]. The 2-D drift-diffusion simulators self-consistently solve the three partial differential equations (4.1)(4.2)(4.3) for the electrostatic potential ψ and for the electron and hole concentrations n and p, respectively [143].

$$\nabla^2(\epsilon\psi) = -q(p - n + N_D^+ - N_A^-) - \rho_s \tag{4.1}$$

$$\frac{\partial n}{\partial t} = \frac{1}{q}\vec{\nabla} \bullet \vec{J_n} - U_n \tag{4.2}$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\vec{\nabla} \bullet \vec{J_p} - U_p \tag{4.3}$$

The electrical behaviour of semiconductor devices is governed by the Poisson's equation given by (4.1) and the continuity equations for electrons and holes given by (4.2) and (4.3) respectively. In equation (4.1), ϵ is the dielectric permittivity; q denotes the electronic charge; N_D^+ and N_A^- represent the ionized impurity concentrations for donors and acceptors, respectively; ρ_s is the surface charge density due to fixed charges in gate insulator or charged interface states. In equations (4.2) and (4.3), J_n and J_p are the current densities for electrons and holes respectively; U_n and U_p are the net recombination rates for electrons and holes respectively.

To obtain the current densities, J_n and J_p, Boltzmann transport theory is applied [124]. By only keeping the lowest-order transport system from the Boltzmann transport equation the following expressions for J_n and J_p are used in the self-consistent calculation.

$$\vec{J_n} = -q\mu_n n\vec{\nabla}\phi_n = -q\mu_n n\vec{\nabla}\psi + qD_n\vec{\nabla}n \tag{4.4}$$

$$\vec{J_p} = -q\mu_p p\vec{\nabla}\phi_p = -q\mu_p p\vec{\nabla}\psi - qD_p\vec{\nabla}p \tag{4.5}$$

where ϕ_n and ϕ_p are the quasi-Fermi potentials for electrons and holes respectively; μ_n and μ_p are the electron and hole mobilities respectively; D_n and D_p are the electron and hole diffusion coefficients respectively and defined by the *Einstein Relations*: $D_{n(p)} = \frac{kT}{q}\mu_{n(p)}$ under the assumption of Maxwell-Boltzmann statistics.

To properly describe the carrier transport within semiconductor devices the drift-diffusion model is not sufficient, especially for carriers under non-equilibrium conditions, as addressed in Chapter 3. For small devices featuring non-equilibrium transport, more advanced approximations of the Boltzmann equation, such as the hydrodynamic transport model, or a statistical solution of the Boltzmann equation, *e.g.* Monte Carlo particle simulation, are required and employed in later chapters.

For the devices under equilibrium condition the drift-diffusion simulation is very efficient and provides adequate descriptions of the electrical behaviour. Provided with proper mobility models for μ_n and μ_p it can account for realistic carrier transport behavior. The mobility models [143] used in this simulation work for the calculations of μ_n and μ_p are given below.

$$\mu_{n(p)} = \frac{\mu_{S,n(p)}}{\left[1 + \left(\frac{\mu_{S,n(p)} E_{||,n(p)}}{v^{sat}_{n(p)}}\right)^{\beta}\right]^{1/\beta}} \tag{4.6}$$

$$\mu_{S,n(p)} = \frac{\mu_{0,n(p)}}{\sqrt{1 + \frac{E_{\perp,n(p)}}{E_{critical}}}} \tag{4.7}$$

Equation (4.6) is the Caughey-Thomas expression which accounts for the carrier heating and the velocity saturation effects due to a high field in the direction of current flow. $v^{sat}_{n(p)}$ is the saturation velocity for electrons (holes); β is a fitting parameter and by default set to 2 for electrons and 1 for holes; $E_{||,n(p)}$ is the electric field component parallel to the electron (hole) current flow; $\mu_{S,n(p)}$ is the low field surface roughness limited electron (hole) mobility and is expressed by equation (4.7). Equation (4.7) is the Perpendicular Electric Field Mobility Model which accounts for the mobility degradation at high perpendicular electric fields due to surface roughness scattering. In equation (4.7), $E_{\perp,n(p)}$ is the electric field component perpendicular to the electron (hole) current flow; $E_{critical}$ is the critical electric field and is a fitting

parameter; $\mu_{0,n(p)}$ is the low field electron (hole) mobility which is either defined as a constant or dependent on the ionized impurity concentrations.

$$\mu_{0,n(p)} = \mu_{min,n(p)} + \frac{\mu_{max,n(p)}\left(\frac{T}{300}\right)^{nu} - \mu_{min,n(p)}}{1+\left(\frac{T}{300}\right)^{xi}\left(\frac{N_{total}}{N_{ref}}\right)^{\alpha_{n,(p)}}} \tag{4.8}$$

Equation (4.8) is an analytical expression of a concentration- and temperature-dependent empirical mobility model which accounts for ionized impurity scattering. T is the temperature; N_{total} is the local total impurity concentration; other parameters are empirical fitting parameters and their default values are taken from [143].

Another important issue for deep submicron device simulation is the concern of quantum mechanical effects. In the 2D device simulators MEDICI and TAURUS many methods have been used to account for the quantum mechanical effects within a confined channel [143,145]. However, the most accurate way is to solve Schrödinger equation self-consistently with Poisson's equation, which has been realized in TAURUS [145]. A 1-dimensional (1D) Poisson-Schrödinger solver by Greg Snider [28] is also used in this project to study the vertical structure of strained Si MODFETs.

4.2 RF Parameter Extraction

Since this chapter studies the strained Si/SiGe MODFETs for RF applications, it is necessary to use some figures of merit to assess the device performance. Using MEDICI transient simulations based on the models described beforehand, small signal analysis is enabled and used to extract RF parameters.

4.2.1 Small signal analysis

In a two-port network, assume $I_1(V_1)$ and $I_2(V_2)$ to be the current (voltage) of the input and output terminal separately, the following equations can be written to describe a Y-parameter network,

$$\begin{cases} I_1 = Y_{11}V_1 + Y_{12}V_2 \\ I_2 = Y_{21}V_1 + Y_{22}V_2 \end{cases} \tag{4.9}$$

where $Y_{ij} = \frac{I_i}{V_j}\Big|_{V_{k,k\neq j}=0}$ (i, j=1 or 2). Let's define a matrix $\widetilde{Y} = \{Y_{ij}\} = G + j\omega C$ to represent these short-circuit admittance parameters, where G is the conductance matrix and C is the capacitance matrix.

Let a small voltage perturbation ΔV_j be applied to contact j at t=0. Using $i_m(t)$ ($v_m(t)$) to denote the total current (voltage) at contact m the small-signal admittance matrix component $\widetilde{Y}$ is given by [146]:

$$\widetilde{Y_{ij}} = \frac{\mathcal{F}\{i_i(t) - I_i(0)\}}{\mathcal{F}\{v_i(t) - V_i(0)\}} = \frac{\mathcal{F}\{i_i(t) - I_i(0)\}}{\Delta V_j} \tag{4.10}$$

where $I_i(0)$ ($V_i(0)$) is the initial current (voltage) at contact i. $\mathcal{F}\{i(t)\}$ is the Fourier transform of the function $i(t)$. Substituting the Fourier decomposition into equation (4.10) and separating it into the frequency dependent and independent parts, the following expressions can be obtained [146]:

$$\begin{aligned} G_{ij} &= \left\{ \frac{I_i(\infty) - I_i(0)}{\Delta V_j} + \frac{\omega}{\Delta V_j} \int_0^\infty [i_i(t) - I_i(\infty)] \sin \omega t dt \right\} \\ C_{ij} &= \left\{ \frac{1}{\Delta V_j} \int_0^\infty [i_i(t) - I_i(\infty)] \sin \omega t dt \right\} \end{aligned} \tag{4.11}$$

where $I_i(\infty)$ gives the final DC value of the current at the contact i.

Scattering parameters (S-parameters) are introduced for microwave transistors to see wave pictures and are normally used in device measurements for frequency analysis. In the two-port network illustrated in Fig. 4.1, the normalized complex wave amplitudes are defined as a_i and b_i, where a_i indicates waves towards the two-port and b_i waves travelling away from it. Then the two-port network can be described by its S-matrix as [147]:

$$\begin{cases} b_1 = S_{11}a_1 + S_{12}a_2 \\ b_2 = S_{21}a_1 + S_{22}a_2 \end{cases} \tag{4.12}$$

The following relationships are used to transform Y-parameters to S-parameters [147]:

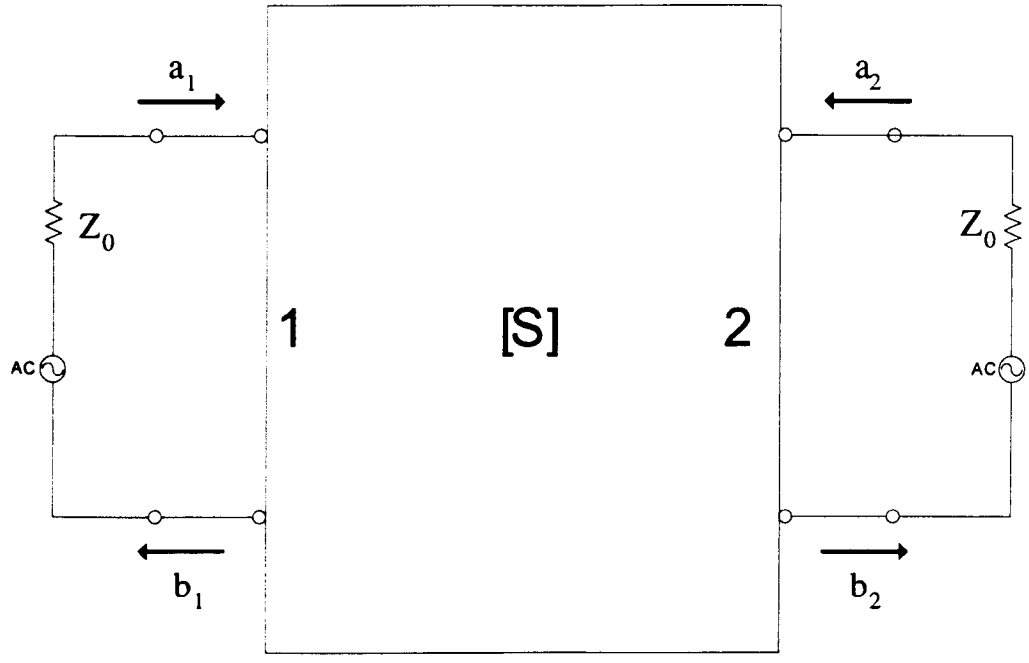


Figure 4.1: A two-port network

$$\begin{cases} S_{11} = \frac{(1-y_{11})(1+y_{22})+y_{12}y_{21}}{(1+y_{11})(1+y_{22})-y_{12}y_{21}} \\ S_{12} = \frac{-2y_{12}}{(1+y_{11})(1+y_{22})-y_{12}y_{21}} \\ S_{21} = \frac{-2y_{21}}{(1+y_{11})(1+y_{22})-y_{12}y_{21}} \\ S_{22} = \frac{(1+y_{11})(1-y_{22})+y_{12}y_{21}}{(1+y_{11})(1+y_{22})-y_{12}y_{21}} \end{cases} \tag{4.13}$$

where $y_{ij} = Y_{ij}Z_0$, Z_0 is the characteristics impedence and Y_{ij} is the Y-parameter component in equation (4.10). The amplitude, S_{ma}, and the phase, S_{ph}, of S-parameters can be expressed by:

$$\begin{aligned} S_{ma,ij} &= |S_{ij}| \\ S_{ph,ij} &= \tan^{-1}\left(\frac{Im(S_{ij})}{Re(S_{ij})}\right) \times 360/2\pi \end{aligned} \tag{4.14}$$

In the MEDICI simulator the Y-parameters can be deduced after a transient simulation with a step change ΔV_g of the gate bias followed by a transient simulation with a step change ΔV_d of the drain bias [148]. The S-parameters can be obtained using the transformations (4.13) from the Y-parameters.

The Field-Effect Transistor (FET) can be treated as a two-port network with a common source terminal. A small-signal equivalent circuit is required to analyze the frequency response of a transistor. Fig. 4.2 is an equivalent circuit [149] which

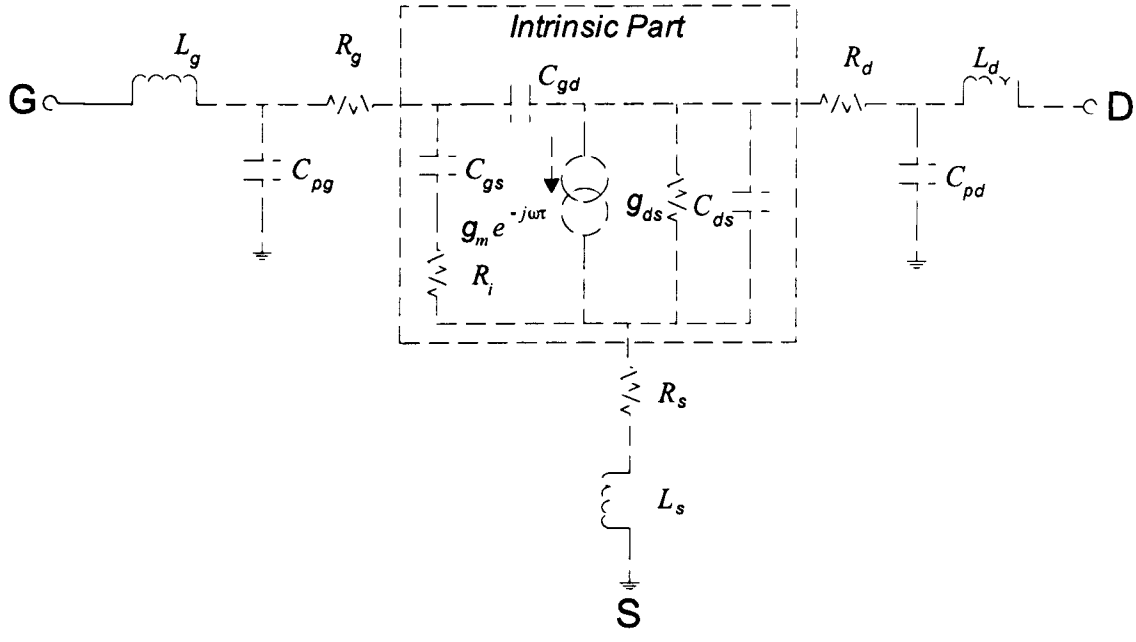


Figure 4.2: Small-signal equivalent circuit of a field effect transistor [149]

contains the external parasitic elements and the intrinsic device including seven unknown parameters .

In Fig. 4.2, L_g, L_d, L_s are the external inductances; C_{pg} and C_{pd} are the external pad capacitances; R_s, R_d and R_g are total terminal resistances including the pad resistance. These external parasitic parameters can be deduced from S-parameter after experimental measurement [150]. The seven intrinsic parameters may be written in terms of the Y parameters: C_{ds}, C_{gd} and C_{gs} are the capacitances between three terminals; R_i is the intrinsic resistance; g_{ds} is the drain transconductance; g_m is the steady-state transconductance and τ is the delay time of the transistor. These parameters can be determined analytically from the following expressions [149]:

$$\begin{cases} C_{gd} = -\frac{\mathrm{Re}(Y_{12})}{\omega} \\ C_{ds} = -\frac{\mathrm{Im}(Y_{22})-\omega C_{gd}}{\omega} \\ C_{gs} = \frac{\mathrm{Im}(Y_{11})-\omega C_{gd}}{\omega}\left(1+\frac{(\mathrm{Re}(Y_{11}))^2}{(\mathrm{Im}(Y_{11})-\omega C_{gd})^2}\right) \\ R_i = \frac{\mathrm{Re}(Y_{11})}{(\mathrm{Im}(Y_{11})-\omega C_{gd})^2+(\mathrm{Re}(Y_{11}))^2} \\ g_{ds} = \mathrm{Re}(Y_{22}) \\ g_m = \sqrt{((\mathrm{Im}(Y_{21})+\omega C_{gd})^2+(\mathrm{Re}(Y_{21}))^2)(1+\omega^2 C_{gs}^2 R_i^2)} \\ \tau = \frac{1}{\omega}\arcsin\left(\frac{-\omega C_{gd}-\mathrm{Im}(Y_{21})-\omega C_{gs}R_i\mathrm{Re}(Y_{21})}{g_m}\right) \end{cases} \tag{4.15}$$

4.2.2 Figures of Merit

The unity current gain frequency, f_T, is defined as the signal input frequency at which the extrapolated small-signal current gain of the transistor equals one. f_T is used as a benchmark to describe the speed of intrinsic devices. The small-signal current gain, G_C, is defined as the amplitude of small-signal drain current to small-signal gate current and is expressed in terms of the extracted Y-parameters by:

$$G_C = \frac{dI_d}{dI_g} = \left[\frac{\partial I_d/\partial V_g}{\partial I_d/\partial V_d}\right] = \left|\frac{Y_{21}}{Y_{11}}\right| \tag{4.16}$$

f_T is then extracted by solving $\log[G_C(\log[f])] = 0$.

After transforming the Y-parameters to S-parameters the maximum frequency of oscillation, f_{max}, can be extracted by solving $\log[G_{MAG}(\log[f])] = 0$, where G_{MAG} is the maximum available gain and is expressed by [151]:

$$G_{MAG} = \left|\frac{S_{21}}{S_{12}}\right| (K - \sqrt{K^2 - 1}),\ if\ K > 1 \tag{4.17}$$

Here, K is the stability factor and is defined by:

$$K = \frac{1 + |S_{11}S_{22} - S_{12}S_{21}|^2 - |S_{11}|^2 - |S_{22}|^2}{2\,|S_{12}S_{21}|} \tag{4.18}$$

When the sum of two sinusoids is applied to the input of a nonlinear device the output contains harmonics of the original frequencies and various intermodulation (IM) products. A classical approximation of PIP3 (Power at the 3rd-order Intercept Point), representing the input signal power leading to excess third-order intermodulation, may be used as a figure of merit for the linearity of investigated devices [152], as illustrated in Fig. 4.3. The analytical expression of PIP3 is obtained by the following deductions.

Assume an input voltage of V applied to a device gives an output current of I. The value of I at an input DC bias V_G with a small AC signal v can be expressed by a Taylor series expansion:

$$I(V_G + v) = I_0 + A_1 v + A_2 v^2 + \cdots \tag{4.19}$$

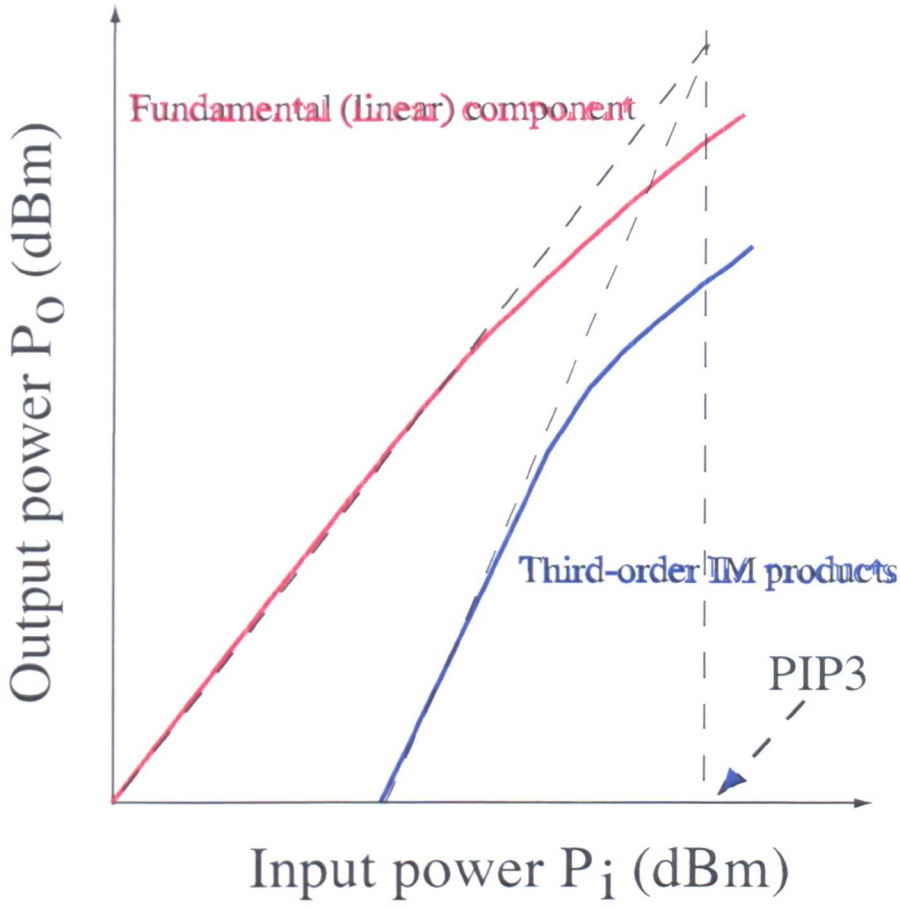


Figure 4.3: Illustration of the definition for PIP3

where $A_n = \frac{1}{n!} g_{m(n-1)}$ and $g_{m(n-1)}$ is the $(n\text{-}1)^{\text{th}}$ derivative of g_m in respect of the gate voltage V_G. Assuming the two input sinusoids with frequencies ω_1 and ω_2 have the same amplitude A, the small signal v can be expressed by $v = A(\cos(\omega_1 t) + \cos(\omega_2 t))$. Substitution of the expression of v into equation (4.19) delivers the output components at frequencies ω_1, ω_2, $2\omega_1 - \omega_2$, $2\omega_2 - \omega_1$ and many other frequencies. The frequency components which are not listed are either far away from the two original frequencies or have small amplitudes and therefore can be ignored in evaluating the linearity upto the third order IM distortion. By only keeping the selected frequency components upto the third order, equation (4.19) becomes:

$$I = I_0 + A_1 A \left(\cos(\omega_1 t) + \cos(\omega_2 t)\right) + \frac{3}{4} A_3 A^3 \left(\cos(2\omega_1 t - \omega_2 t) + \cos(2\omega_2 t - \omega_1 t)\right) \tag{4.20}$$

To have the third-order intercept point illustrated in Fig. 4.3 for PIP3, the amplitude of the linear term and that of the third-order term in equation (4.19) is equalled giving $A_1 A = \frac{3}{4} A_3 A^3$. Thus, assuming the input resistance to be R_s, the PIP3 can be expressed by:

$$PIP3 = \frac{A^2}{2R_s} = \frac{2A_1}{3A_3 R_s} = \frac{4g_m}{g_{m2} R_s} \tag{4.21}$$

where g_m is the transconductance, g_{m2} is the second derivative of g_m in respect of the gate voltage V_G, and R_s=50 Ω is the load resistance.

4.3 Device Calibration

4.3.1 Calibration Methodology

It is unlikely to run device simulations with default parameters and obtain good agreements with experimental characteristics. The default parameters used in commercial simulators often depend on processing conditions and specific models used. Therefore, generic parameters which can fully describe the internal behaviours of all devices are not available. The role of device calibration is to reproduce experimental device characteristics by adjusting a series of parameters in a manner with physical meaning. Before the calibration it is necessary to obtain as much as possible detailed information of studied device. The device structure, the dimensions in both directions, the doping profiles and their activation rates, the contact resistances, the interface states and the device measurements are essential.

The calibration using MEDICI starts with the construction of the MEDICI input file to describe the device structure illustrated in Fig. 4.4 and include all necessary material parameters. The device description should contain the device structure and dimensions, the doping profiles and the mesh definitions. The material parameters, including the bandgap, the conduction and valence band offsets, the effective masses, the density of states and the dielectric permittivity relevant to the Si/SiGe heterostructure are calculated according to the theory addressed in Chapter 3. Appropriate mobility models are then selected from the available model list of MEDICI

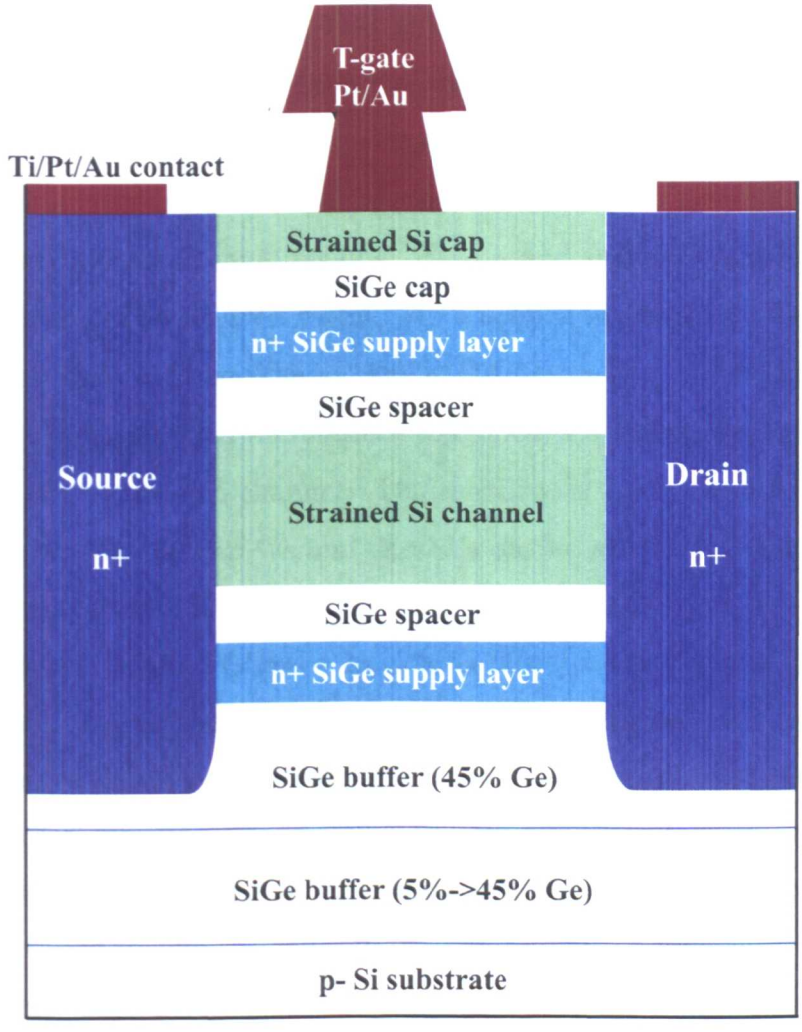


Figure 4.4: Illustration of simulated strained Si/SiGe MODFET

and applied to specific regions of the device. The Caughy-Thomas lateral field dependent and the perpendicular field dependent mobility models [143] may be used to describe the transport within Si/SiGe MODFETs. The Caughy-Thomas lateral field dependent mobility model, which accounts for the high field velocity saturation effects, has two important parameters, the saturation velocity v_{sat} and the fitting parameter β given by equation (4.6). β is usually set to 2 for electrons and 1 for holes. The saturation velocity v_{sat} for electrons in Si is by default set to 0.97×10^7 cm/s, while that in strained Si may be 10-20% higher, *i.e.*, about $1.1\text{-}1.2\times10^7$ cm/s [153], due to the enhanced transport in strained Si.

The perpendicular field dependent mobility model, described by equation (4.7), has two parameters, the vertical critical field $E_{critical}$ and the low field mobility μ,

if it is assumed as a constant. For the device structure studied in this chapter, as illustrated in Fig. 4.5, the source-drain distance is much larger than the gate length, so that different positions in the channel may have different high perpendicular field effects. Thus it is reasonable to divide the channel into several parts and use different $E_{critical}$ for the different channel regions. $E_{critical}$ and μ should also be different in the different layers. The electron mobility in a confined strained Si 2DEG has been reported to be 1000-2900 cm^2/Vs at 300 K [46, 51 53, 154]. The mobility in the strained Si cap layer suffers from interface roughness scattering and should be lower than that in the strained Si channel. The mobility in the relaxed SiGe layer has been reported at 100-800 cm^2/Vs and depends on the germanium content and the impurity concentration [155].

The calibration is then followed by a loop of adjusting a series of parameters discussed above. The first step is the adjustment of $E_{critical}$ and μ to obtain a agreement between the simulated and measured I_D-V_G characteristics at low drain bias, *e.g.*, 50 mV or 100 mV. The second step is to tune $E_{critical}$ and v_{sat} in order to reproduce the I_D-V_G characteristics at high drain voltage. Contact resistances may be included and varied to help the calibration. The calibration may need to come back to the parameter adjustment at low drain voltage if it fails to agree with the experimental data at high drain bias. Activation rates of the doping concentrations in the supply layers may be changed according to the information from the fabrication to obtain different sheet carrier density in the channel. Successful calibrations at both low and high drain biases then lead to the calibration of I_D-V_D characteristics for a whole range of gate biases. Care need to be taken on all parameters at this stage, especially v_{sat}. These three steps may need be repeated in order to obtain a global fitting with physical parameters. However, if the simulation can't fit the high gate I_D-V_D characteristics after several loops, the velocity overshoot may be the cause and a more advanced simulation technique, such as the hydrodynamic model, is required.

4.3.2 Device Structure and Calibration Results

The strained Si/SiGe MODFETs studied in this project are based on a 0.25

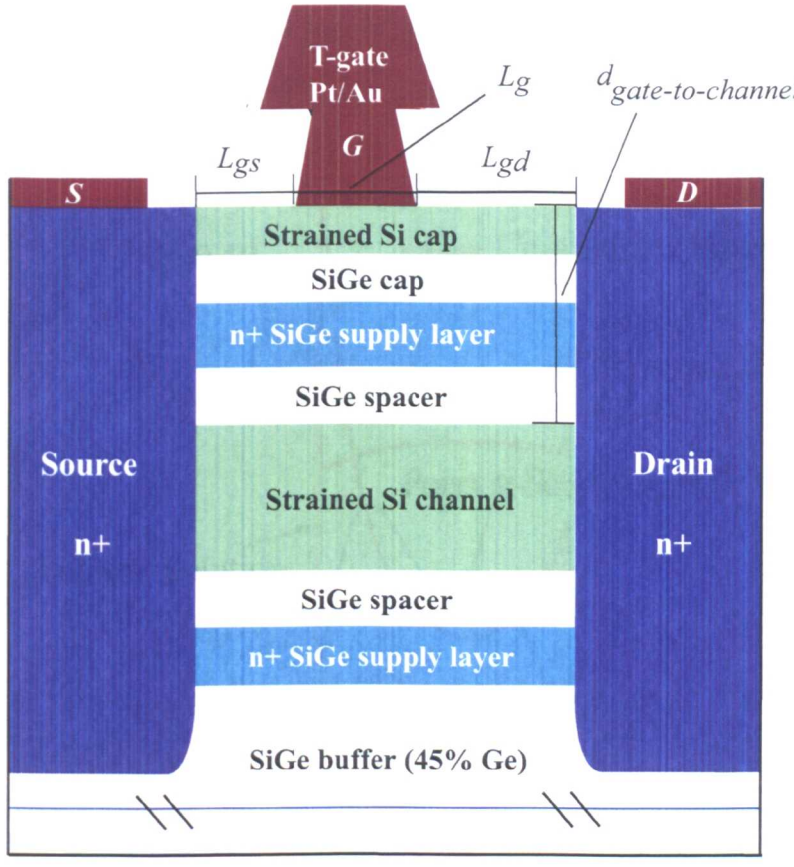


Figure 4.5: Schematic of strained Si/SiGe MODFET

μm and a 70 nm n-type buried strained Si channel Si/SiGe MODFETs fabricated by Daimler Chrysler, as illustrated in Fig. 4.4. The layer sequence of the 0.25 μm MODFET is [30]: a p^- substrate with ρ>1000 Ω·cm; a relaxed SiGe buffer with linearly graded Ge content from 5% to 45%; a 4 nm SiGe supply layer with a dopant concentration of N_D=4×10^{18} cm^{-3}; a 3 nm SiGe spacer; a 9 nm strained Si channel; a 3 nm SiGe spacer; a 3.5 nm SiGe supply layer with a doping level of N_D=1.5×10^{19} cm^{-3}; a 3 nm SiGe cap layer and a 4 nm Si cap layer. The T-shape Au/Pt gate is located asymmetrically with source-gate distance L_{gs}=0.5 μm while total drain-source distance L_{ds}=1.5 μm, as illustrated in Fig. 4.5.

Fig. 4.6 shows the 1D band diagram and the comparisons of the electron distributions by the 1D Greg-Snider solver based on the solutions of Poisson's equation and Poisson-Schrödinger. From the 1D band diagram it is evident that there are parasitic conduction paths coming with the 2-dimensional electron gas. The confined electron distribution in the channel is obtained from the 1D Poisson-Schrödinger

solution (only solving the channel region) and compared with the solution of the 1D Poisson equation in Fig. 4.6.

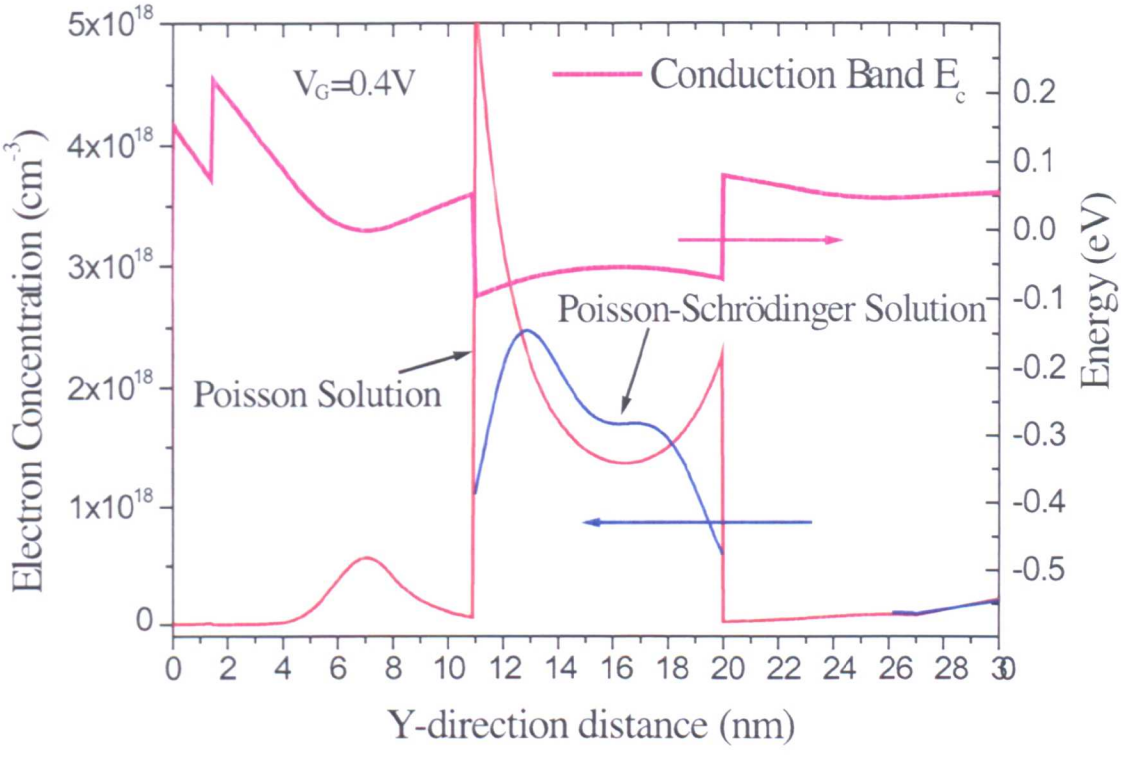


Figure 4.6: 1D band diagram and electron distributions from the 1D Greg-Snider solver based on the solutions of Poisson's equation and Poisson-Schrödinger

Followed the methodology addressed, a successful calibration is carried out in respect of the 0.25 μm MODFET using MEDICI. Figs. 4.7 and 4.8 show the I_D-V_G and I_D-V_D characteristics (simulated and experimental) respectively and demonstrate the good agreement between measurements and simulations. The calibrated low field mobility in the 2-DEG is 1500 cm^2/Vs which is in agreement with the value suggested by Daimler Chrysler [156]. However, slight discrepancies are observed at high V_G and V_D which may be attributable to self-heating [61]. The simulations predict a current gain cutoff frequency, f_T, of about 24 GHz, as indicated in Fig. 4.9, also in agreement with experimental data. However, leakage in the buffer in this device causes the on current and off current ratio I_{on}/I_{off} to be less than 10.

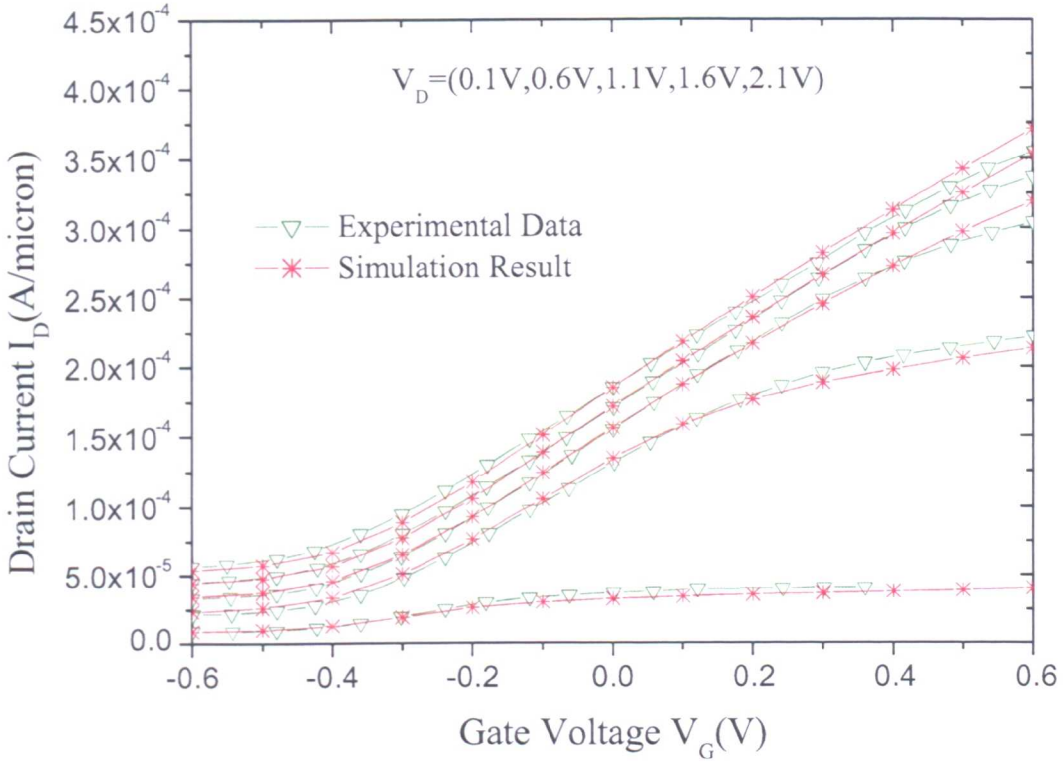


Figure 4.7: Calibrated I_D-V_G characteristics of the 0.25 μm MODFET

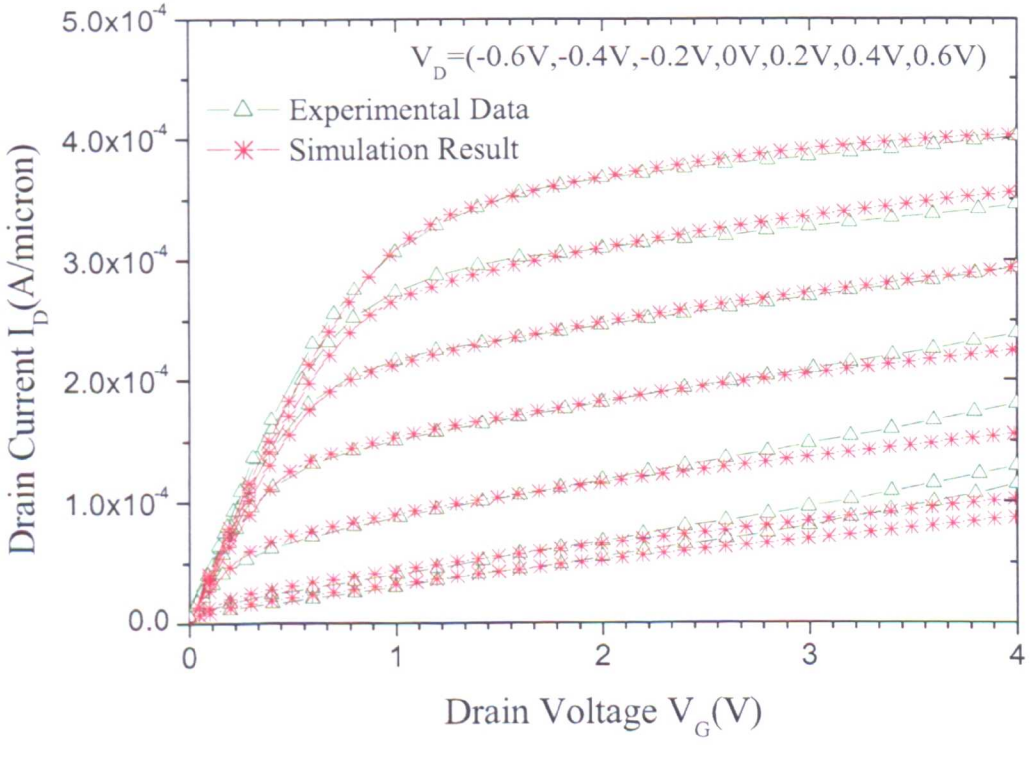


Figure 4.8: Calibrated I_D-V_D characteristics of the 0.25 μm MODFET

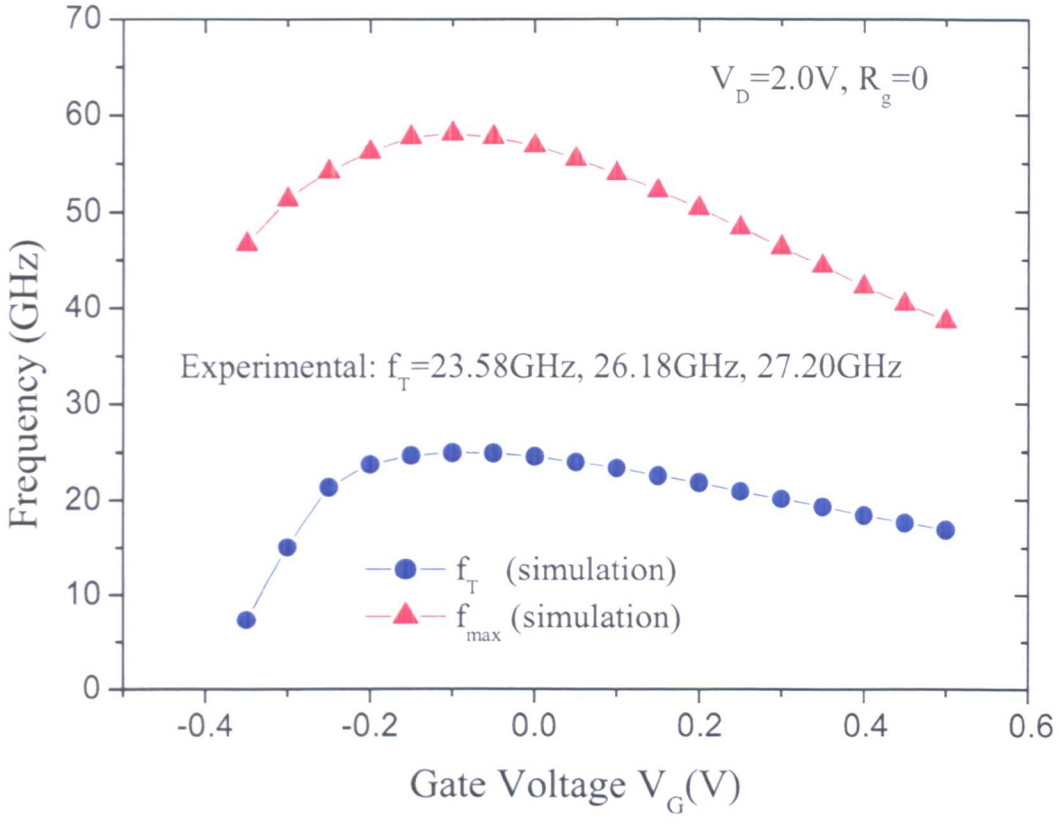


Figure 4.9: Extracted f_T and f_{max} characteristics of the 0.25 μm MODFET

A comprehensive calibration in respect of a 70 nm n-type Si/SiGe MODFET fabricated by Daimler Chrysler is also carried out. The layer sequence (from bottom to top) of the calibrated device is [33]: a p^- substrate with ρ>1000 Ω·cm; a relaxed SiGe buffer with linearly graded Ge content from 5% to 45%; a 5 nm SiGe supply layer with a dopant concentration of 2.5×10^{18} cm^{-3}; a 3.5 nm SiGe spacer; a 9 nm strained Si channel; a 3 nm SiGe spacer; a 5 nm SiGe supply layer with a doping level of 1.0×10^{19} cm^{-3}; a 6 nm SiGe cap layer and a 2 nm Si cap layer. The T-shape Au/Pt gate is located with a source-gate distance of L_{gs}=0.5 μm while the total drain-source distance is L_{ds}=1.0 μm. The main differences of the 70 nm device compared to the 0.25 μm are the reduced gate length, the reduced drain-source distance and the shifted gate (increased ratio of L_{gs}/L_{ds}) aiming for high RF performance, and the increased gate-to-channel distance aiming for high linearity.

In the calibration of the 70 nm MODFET, the channel has been divided into three regions in which different $E_{critical}$ have been applied. The calibrated low field

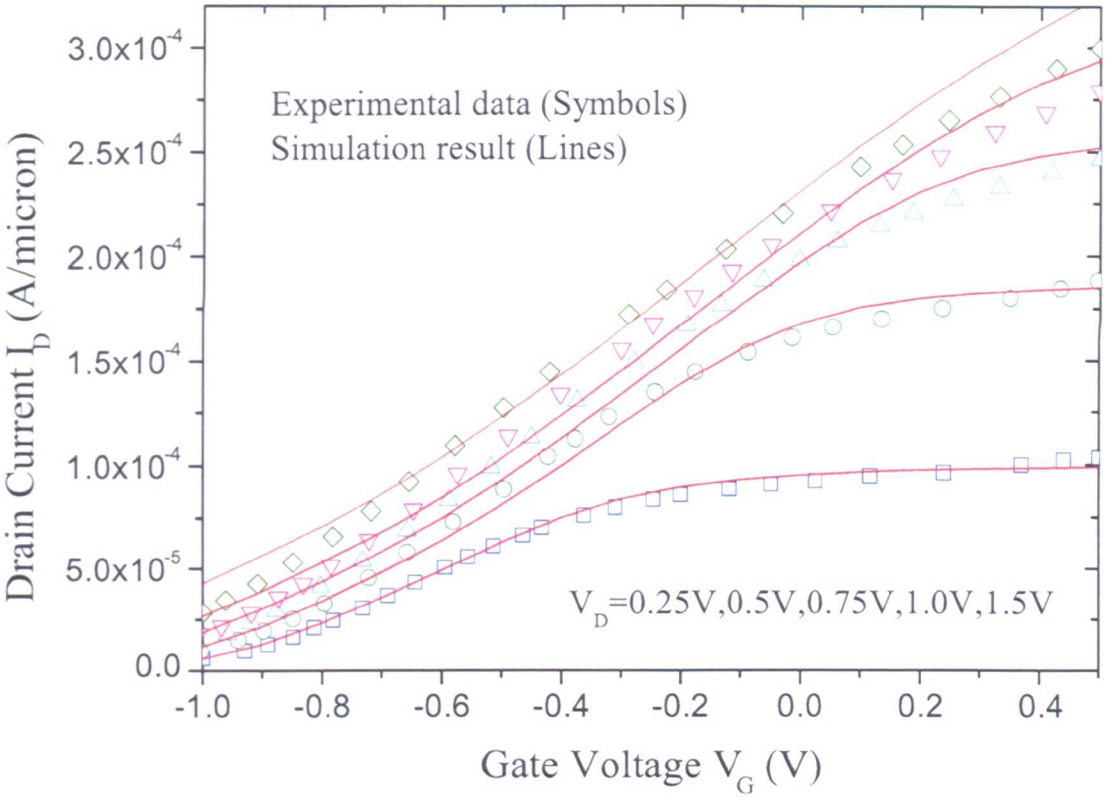


Figure 4.10: Calibrated I_D-V_G characteristics of the 70 nm MODFET

mobility in the 2-DEG is 1600 cm^2/Vs. Fig. 4.10 shows the I_D-V_G characteristics (simulated and experimental) and demonstrates the good agreement between measurements and simulations. The slight discrepancies at high V_D can be again attributed to self-heating effects which are not included in the simulations [61].

4.4 Impact of the Device Geometry on RF performance

4.4.1 An Qualitative Analysis

Before using MEDICI to simulate the impact of the device geometry, it is essential to qualitatively analyze the importance of different parameters in the complicated device structure of MODFET. Therefore, instead of extracting the cutoff frequency, f_T, and the maximum available frequency, f_{max}, from equation (4.16) and (4.17) for MEDICI simulations, analytical expressions for f_T and f_{max} are used to qualitatively

analyze the dependence on the RF performance. The f_T and f_{max} are functions of the seven intrinsic parameters (see equation (4.15)) and the external pad resistances R_g, R_s and R_d. They can be calculated using the following relations [157,158]:

$$f_T = \frac{g_m/2\pi}{(C_{gs}+C_{gd})[1+(R_s+R_d)g_{ds}]+C_{gd}g_m(R_s+R_d)} \tag{4.22}$$

$$f_{max} = \frac{f_T}{2\sqrt{2\pi f_T R_g C_{gd}+(R_g+R_i)g_{ds}}} \tag{4.23}$$

From equations (4.22) and (4.23), in order to obtain high f_T and f_{max}, g_m has to be as large as possible; C_{gs}, C_{gd}, C_{ds}, R_i, R_g R_s, R_d and g_{ds} need to be as small as possible. These parameters are sensitive to both lateral and vertical device designs.

In the lateral direction, the source-to-drain distance, L_{ds}, the gate position, L_{gs}/L_{ds}, and the physical gate length, L_g, are important. Increasing L_{ds} leads to small capacitances but on the other hand it induces more intrinsic resistance R_i. Different L_{gs}/L_{ds} causes the change of C_{gs}/C_{gd} which results in different f_{max}. Shrinking L_g is a commonly used approach to increase the device speed, however, the g_{ds} increase induced by the 2-D effects degrades the device performance. Therefore, careful trade-off designs of these lateral dimensions are required in order to obtain a high RF performance.

In the vertical direction, the layer structure and the doping profiles need to be optimized in order to achieve high carrier density in the channel. In order to analyze the RF performance dependence of the device design in the vertical direction, the following deduction may be used to obtain another expression for f_T to evaluate the RF performance. Neglecting the effects of the parasitic resistance R_s and R_d, the expression for f_T (4.22) may be simplified to [159]:

$$f_T = \frac{g_m}{2\pi(C_{gs}+C_{gd})} = \frac{g_m}{2\pi c_g W L_{eff}} \tag{4.24}$$

where L_{eff} is the effective gate length which depends on L_g, L_{ds} and L_{gs}; W is the device width; c_g is the effective gate capacitance per unit area between the gate and the 2-DEG channel and may be expressed as $c_g = \frac{\epsilon}{d+\Delta d}$ (ϵ is the effective dielectric constant of the layers above the 2-DEG); d is the distance from the gate to the

2-DEG interface and Δd is the effective distance between the heterojunction and the 2-DEG. When increasing the gate-to-channel separation d, the gate capacitance c_g decreases. However, the increase in the gate-to-channel separation results in a losing of the gate control over the channel and induces a lower g_m and therefore degrades f_T. To evaluate the gate control on the 2-DEG channel the modulation efficiency (ME), as proposed by Foisy *et al* [160], can be used:

$$ME = \frac{\partial n_s / \partial V_G}{\partial (n_s + n_p) / \partial V_G} \tag{4.25}$$

where n_s is the sheet carrier density in the channel and n_p is the parasitic sheet carrier density in the layers above the channel which limits the ME to be less than 1. A high ME means a good control of the gate over the channel and leads to a high transconductance, g_m. It is obvious that the modulation efficiency at a fixed gate voltage depends on the thickness of each layer and the doping concentrations in the supply layers. We can now reformulate analytical expression for f_T (see equation (4.24)) using the definition of the ME as [160]:

$$f_T = \frac{v_{eff}}{2\pi L_{eff}} = \frac{v_{sat}}{2\pi L_{eff}} ME \tag{4.26}$$

where v_{eff} and v_{sat} denote the effective and peak velocities in the 2-DEG channel, respectively.

It can be seen from equations (4.22)-(4.26) that L_g, L_{ds}, L_{gs}, the thickness of each layer and the concentration in each supply layer play important roles in determining the RF performance and careful design of these parameters are required. The effects of these parameters on the device performance are studied by numerical simulations in next section. The effects from the external resistances and capacitances also need to be taken into account to evaluate the real RF performance but they are not the focus of this thesis and will be neglected in the following work.

4.4.2 Simulation Results

Followed the qualitative analysis made in last section, this section uses the numerical simulator MEDICI to study the effects of different device designs, *i.e.*, L_g, L_{ds}

and L_{gs} in the lateral direction, and the thicknesses of each layer and the concentrations in the supply layers in the vertical direction. The simulations carried out in this section are based on the calibration in respect of the 0.25 μm n-type Daimler Chrysler strained Si/SiGe MODFET discussed in section 4.3.2.

4.4.2.1 Impact of the lateral dimensions

Since the speed of FET is intrinsically limited by the electron transit time, the most obvious approach to improve the device speed is to reduce the gate length, as described by equation (4.26). However, the 2-D effects due to the gate length scaling (*e.g.* DIBL) affect the threshold voltage and subthreshold slope and increase the off-state current (see Fig. 4.11). When the gate length is aggressively scaled the gate begins to lose control over the channel and the parasitic conduction layers, which results in a saturation, possibly in a reduction of transconductance g_m and in an increase of drain conductance g_{ds}. These will decrease the cutoff frequency f_T and the voltage gain G (defined as $V_{out}/V_{in} = g_m/g_{ds}$).

The source-drain distance L_{ds} is crucial in reducing the intrinsic resistance R_i. However, reducing L_{ds} also decreases the voltage gain G due to increased drain conductance g_{ds}. It has been suggested that raising the source/drain junctions may compensate for this [55] effect. Scaling L_{gs} and L_{ds} also increases C_{gs} and C_{gd} and induces a second pronounced peak in the g_m characteristics (see Fig. 4.12). This is due to the increase in the current density within the Si cap layer between the source/drain and the gate at high V_G when scaling L_{gs} and L_{ds} [53].

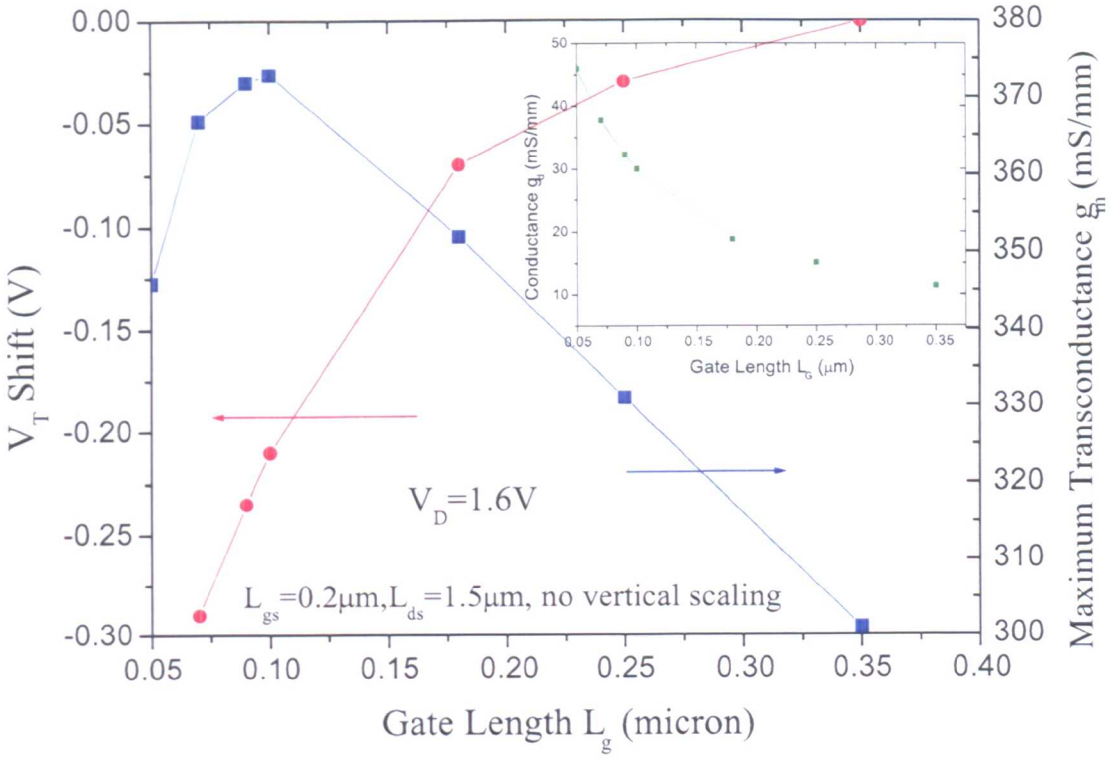


Figure 4.11: Threshold voltage shift and transconductance characteristics versus the gate length; the inset is the output conductance characteristics

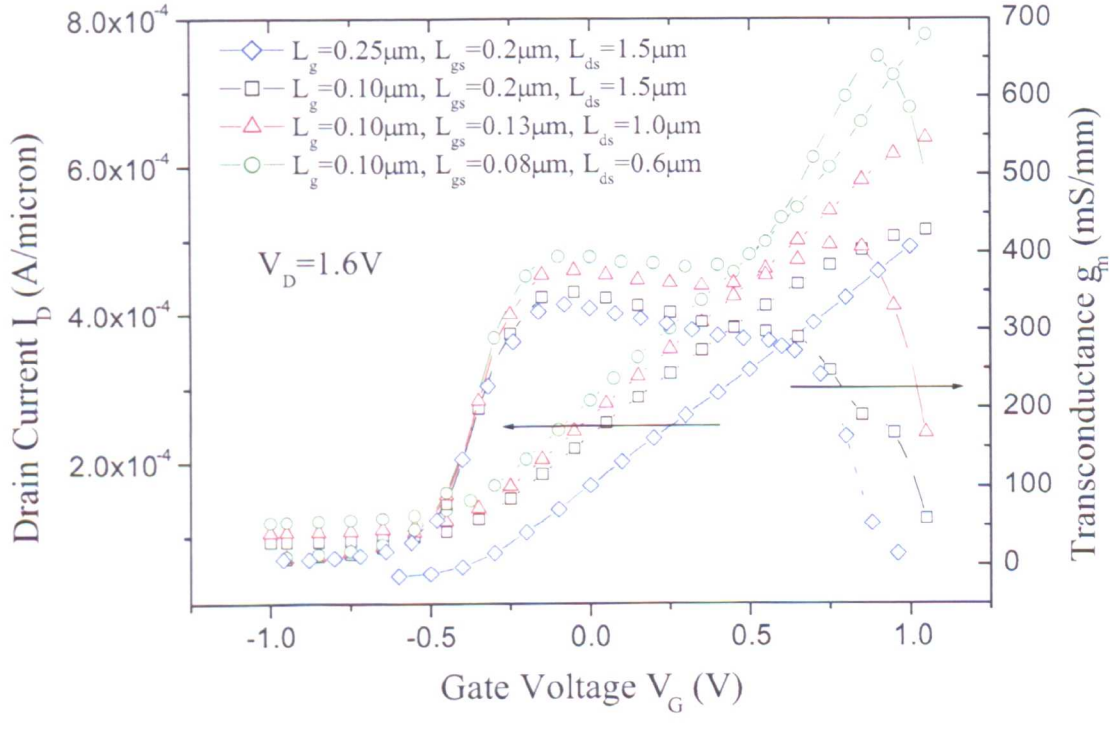


Figure 4.12: Drain current and transconductance characteristics with different lateral scaling strategies

Figs. 4.13 and 4.14 show the effects of changing the L_{gs}/L_{ds} ratio and the source-drain distance L_{ds}. For a given L_{ds}, decreasing the L_{gs}/L_{ds} ratio increases the C_{gs}/C_{gd} ratio, which will generally tend to increase f_{max}. However, the decrease in L_{gs}/L_{ds} also tends to increase the output (drain) conductance, g_{ds}. This increase in g_{ds} will reduce f_{max} by decreasing the effect of C_{gs}/C_{gd} on f_{max}. This effect is particular evident for the layer structures considered in this work due to their relatively high g_{ds}. The RF characteristics as a function of the gate length scaling, with a L_{gs}/L_{ds} ratio 0.5 μm/1.5 μm, are shown in Fig. 4.15. The RF performance improvements with different saturation velocities (also shown in Fig. 4.15) indicate that the effect of velocity overshoot may also lead to an improvement in the RF performance.

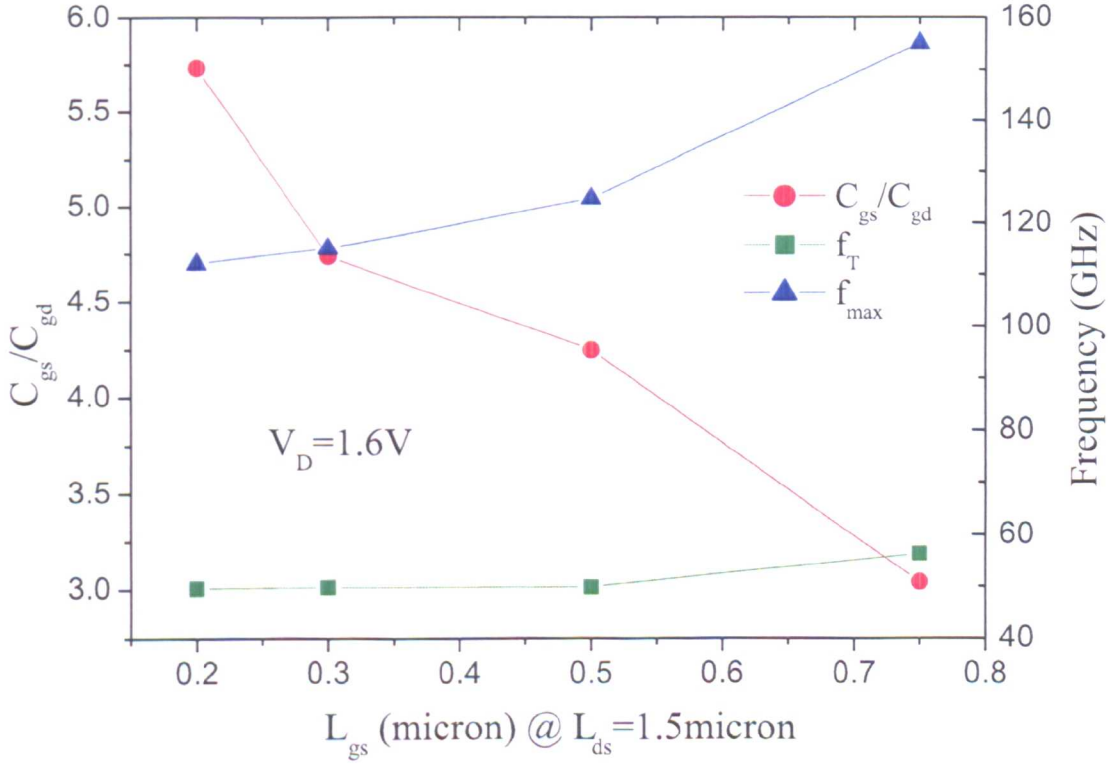


Figure 4.13: Effects of different L_{gs}/L_{ds} ratio (different L_{gs} with a fixed L_{ds}) on the device characteristics (L_g=0.1 μm; all values are extracted from the MEDICI transient simulation without R_g, R_d and R_s)

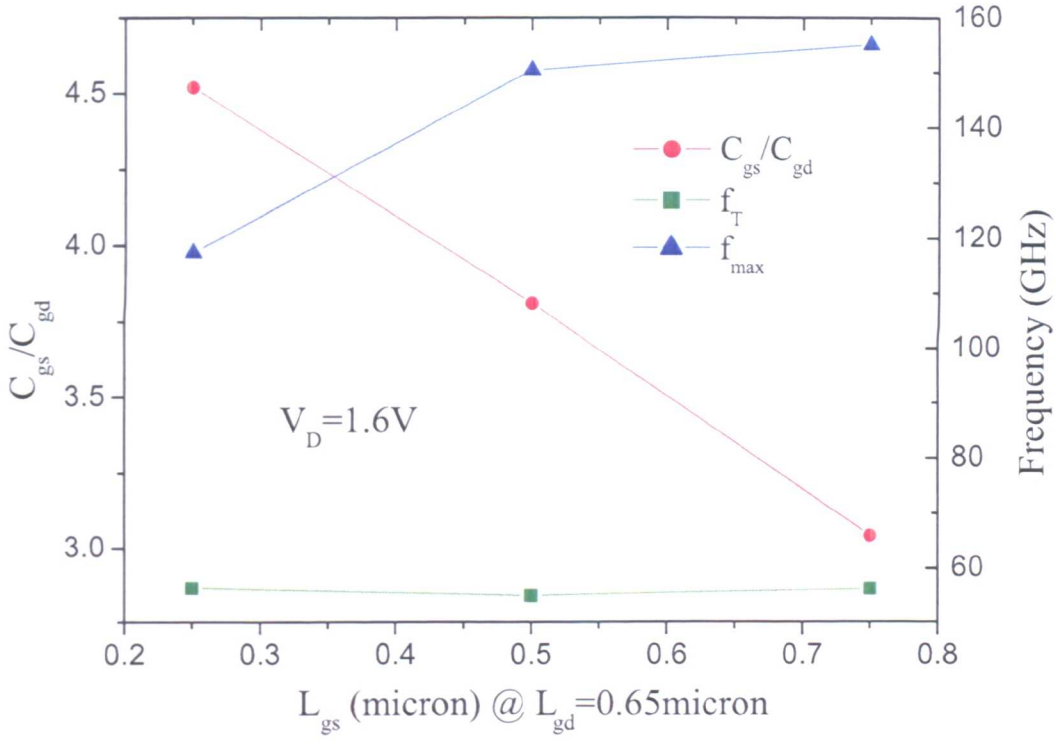


Figure 4.14: Effects of different L_{gs}/L_{ds} ratio (different L_{gs} with a fixed L_{gd}) on the device characteristics (L_g=0.1 μm; all values are extracted from the MEDICI transient simulation without R_g, R_d and R_s)

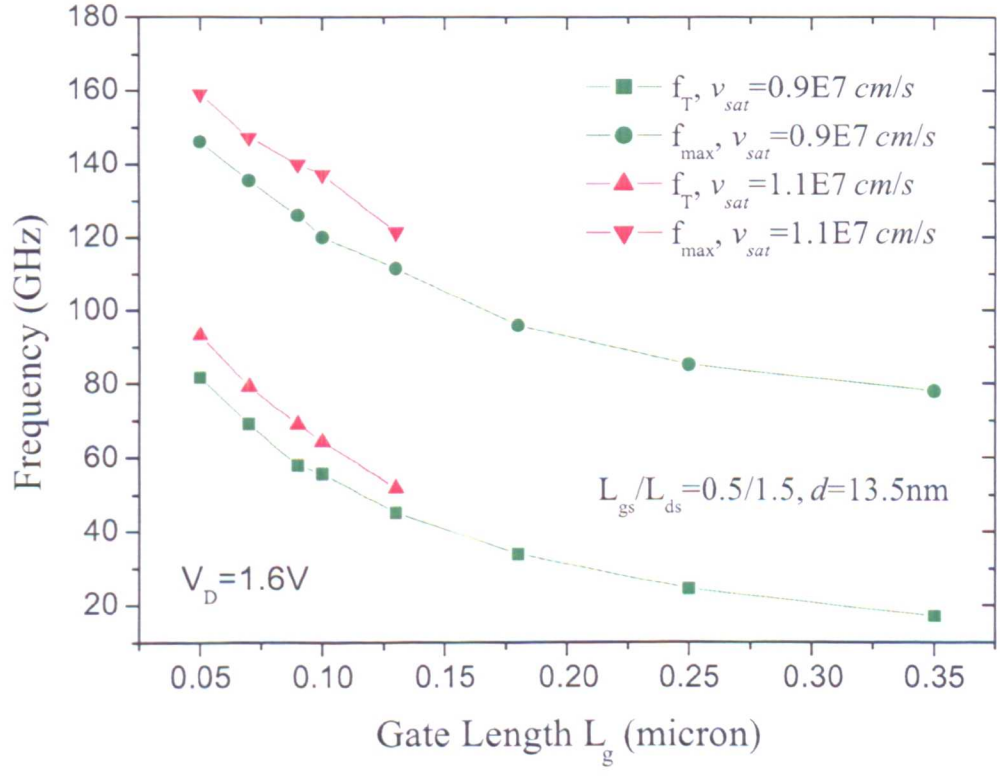


Figure 4.15: Intrinsic RF characteristics versus the gate length at V_D=1.6 V (the gate-to-channel distance=13.5 nm)

4.4.2.2 Effects of different vertical structures

Optimization of the vertical structure may help to improve the device performance and partially suppress the 2-D effects due to the gate length scaling. In the case of the MODFET structure studied in this chapter (Fig. 4.5), the key parameters associated with its vertical structure are the thickness of each layer and the doping concentrations in the supply layers. The thickness of the 2-DEG strained Si channel is a sensitive parameter for the carrier confinement of the 2-DEG.

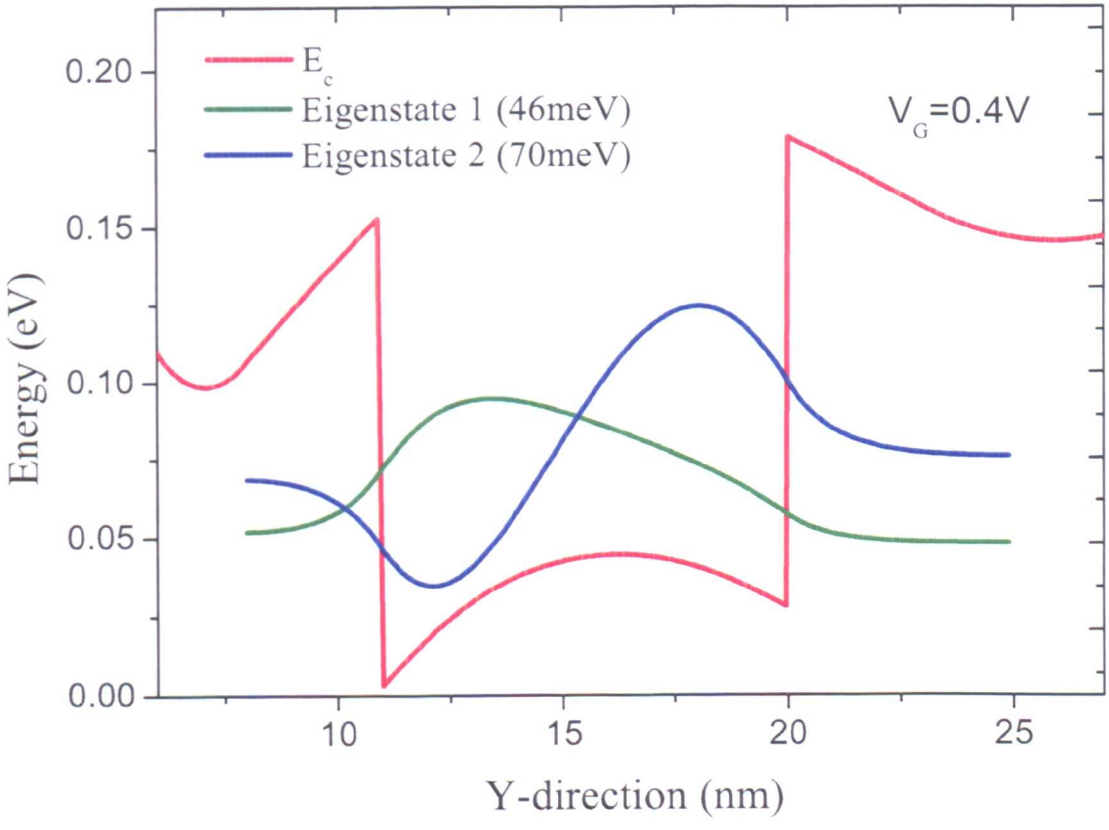
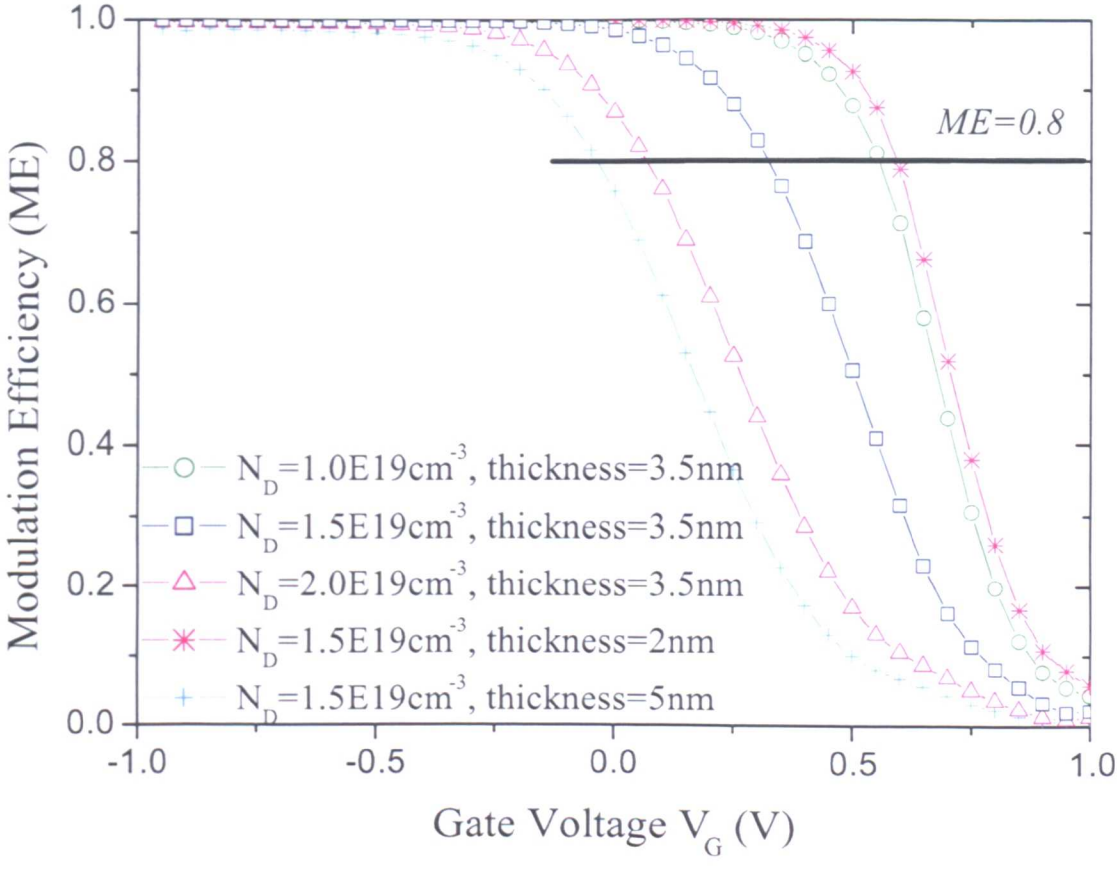


Figure 4.16: The conduction band diagram and the wave functions in a 9 nm thick strained Si channel

Fig. 4.16 illustrates the 1D (vertical direction) conduction band profile and the wave functions within a 9nm thick strained Si channel by 1D Poisson-Schrödinger solutions. Due to quantum confinement, the peak electron concentration appears close to the middle of the channel and the shape of the electron distribution and

the sheet carrier density within the channel depends on the width and the depth of the quantum well. The depth of the well may be varied by changing the Ge content within SiGe layers but too high Ge content causes the problems of layer growth. The width of the quantum well, *i.e.*, the thickness of the channel, needs to be properly chosen. Although the peak concentration increases with decreasing channel thickness, the sheet carrier density in the channel reduces. On the other hand, a too large channel thickness causes carriers to move away from the gate which increases output conductance g_{ds}. The channel thickness is also limited by the critical thickness which decreases as the Ge content in the adjacent SiGe layers increases.

The aim of the optimization in the vertical structure is to maximize the sheet carrier density in the channel and maintain a good gate control over the channel. *ME* defined by equation (4.25) is appropriate to evaluate the gate control on the channel. The gate-to-channel separation requires a careful consideration in order to suppress the SCE introduced by the gate length scaling. The gate-to-channel separation is sum of the thicknesses of the Si and SiGe cap layers, the SiGe doping layer and the SiGe spacer layer. Reducing the Si and the SiGe cap layer thicknesses improves the control of the gate on the device. However, reducing the SiGe cap layer thickness leads to an increase in parasitic conduction within the Si cap. Shrinking the SiGe spacer layer increases the *ME* and the sheet carrier density in the channel but also reduces the mobility within the channel due to an increase of remote ionized impurity scattering. Varying the thickness of the doping layer and their doping concentration changes [161] the capability of supplying carriers into the channel.



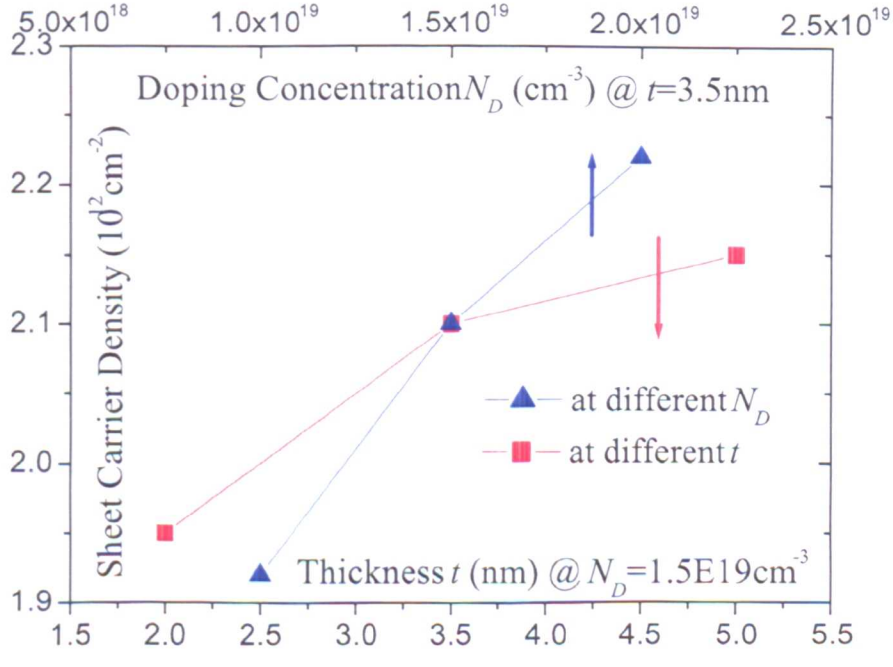


Figure 4.17: Effects of the front supply layer thickness and the doping concentration on the modulation efficiency and the sheet carrier density in the channel (the sheet carrier densities plotted in the second figure are extracted at the ME=0.8 within the first figure)

Fig. 4.17 illustrates the effects of the thickness and the doping concentration in the front SiGe doping supply layer on the modulation efficiency which was obtained from the 2D simulator MEDICI. To illustrate the effects on the carrier density in the channel the sheet carrier densities are also extracted from Fig. 4.17 at ME=0.8 (as an example) and plotted in the same figure as a function of the layer thickness and the doping level. It can be seen from Fig. 4.17 that reducing the doping concentration in the front doping supply layer or decreasing the thickness of this layer decreases the sheet carrier density in the channel and causes a slightly sharper fall off of the ME as the gate voltage is increased. This leads to a reduced gate operation range. However, a too high doping concentration may induce severe mobility reduction in the channel and possibly increase the parasitic carrier density in the layers above the channel.

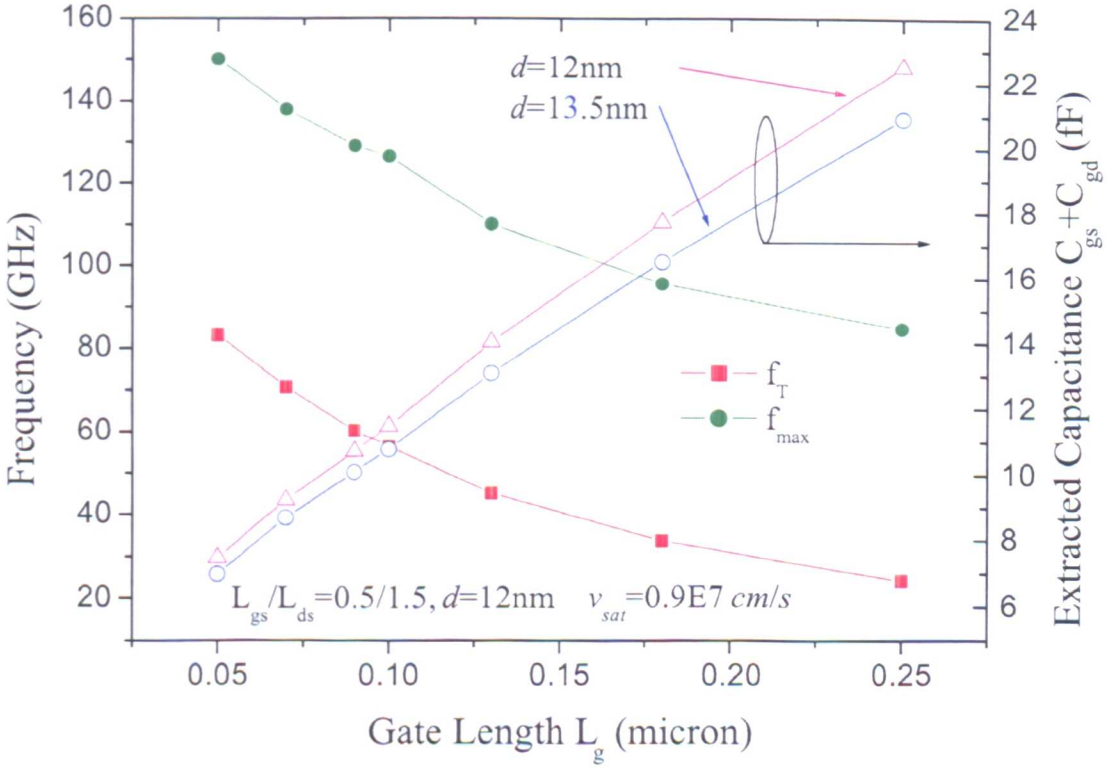


Figure 4.18: Extracted intrinsic RF characteristics and capacitance ($C_{gs} + C_{gd}$) versus the gate length at V_D=1.6 V, the gate-to-channel distance=12 nm

The reduction in g_m (see Fig. 4.11) limits f_T during the gate length scaling. However, this negative effect may be partially offseted by the reduction of the gate-

to-channel separation, d [55]. The RF characteristics versus gate length with a reduced d is presented in Fig. 4.18. The reduction of the d decreases the intrinsic capacitance (the sum of C_{gs} and C_{gd}) and results in an improved RF performance when compared to RF performance in Fig. 4.15.

It is therefore necessary to optimize the both lateral and vertical architectures in order to achieve RF performance improvements in sub-100 nm devices. Nevertheless, parasitic effects become increasingly important with decreasing gate lengths, which together with the 2-D effects, limit improvement of the device performance.

4.5 Impact of Different Device Designs on Linearity

The optimization of Si/SiGe MODFET device geometry is necessary in order to achieve an improved performance by having better quantum confinement and high modulation efficiency which help to achieve high density and high mobility of carriers in the channel. However, the existence of low-mobility parasitic conduction paths in these devices limits device performance and range of operation. At high current levels, the carrier density in the low mobility slab doping layers above the channel increases, screening further modulation of the channel carrier concentration, which in turn limits device linearity and also degrades device performance.

This section discusses the impact of different device designs on linearity. The simulations are based on the calibration in respect of the 70 nm n-type Daimler Chrysler strained Si/SiGe MODFET given in section 4.3.2. The figure of merit for linearity, PIP3, defined by equation (4.21), is used. A larger PIP3 means better linearity, *i.e.*, a wide device operation range featuring small intermodulation distortions.

4.5.1 Impact of the Device Geometry

The gate-to-channel distance, d, is the key parameter which affects the gate control on the conduction layers. The decrease of d helps to achieve a high transconductance and better RF performance. However, small gate-to-channel separations degrade device linearity by compressing the transconductance characteristics. As shown in

Fig. 4.19, the devices with small gate-to-channel separations have a lower PIP3 representing worse linearity performance. However their larger transconductances result in higher RF performance as discussed in section 4.4.

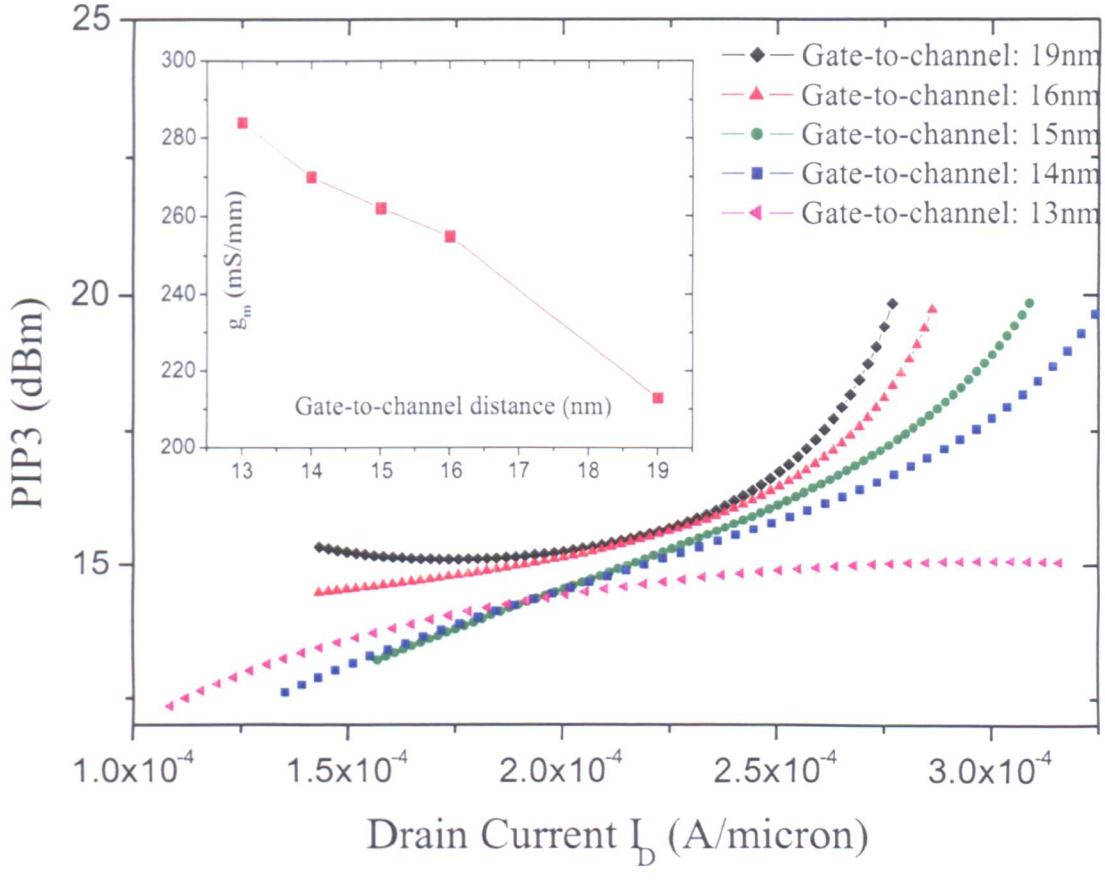


Figure 4.19: The device PIP3 and transconductance g_m (the inset) characteristics with different gate-to-channel distances

The lateral device design also affects linearity. The source-drain distance is larger than the physical gate length which helps to reduce the gate-to-contact parasitic capacitances and increase the breakdown voltage. The transconductance is sensitive to the changes of lateral dimensions as they affect the series resistance. The source series resistance changes the shape of the I_D-V_G characteristics and therefore affects linearity. Fig. 4.20 shows that although increasing the source-gate separation degrades the transconductance and drive current, it flattens the transconductance characteristics and does help to improve the linearity. The reduction of g_m due to the large L_{gs} degrades RF performance, as shown in Fig. 4.14 in section 4.4.

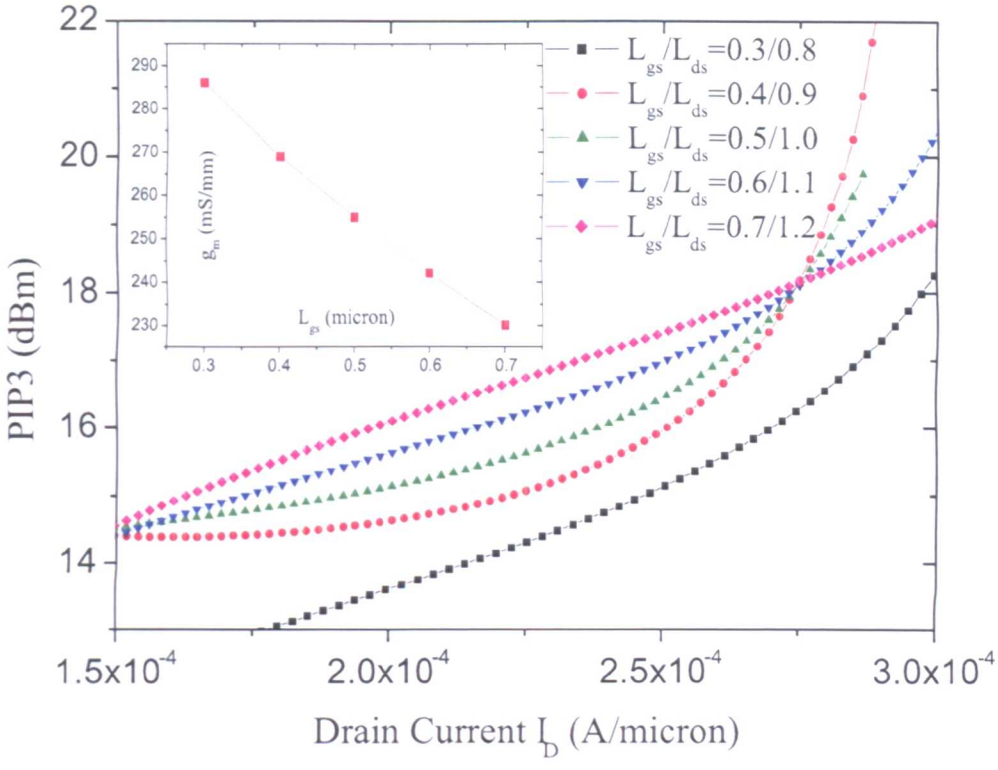


Figure 4.20: The device PIP3 and transconductance g_m (the inset) characteristics with different gate-to-source distances L_{gs}

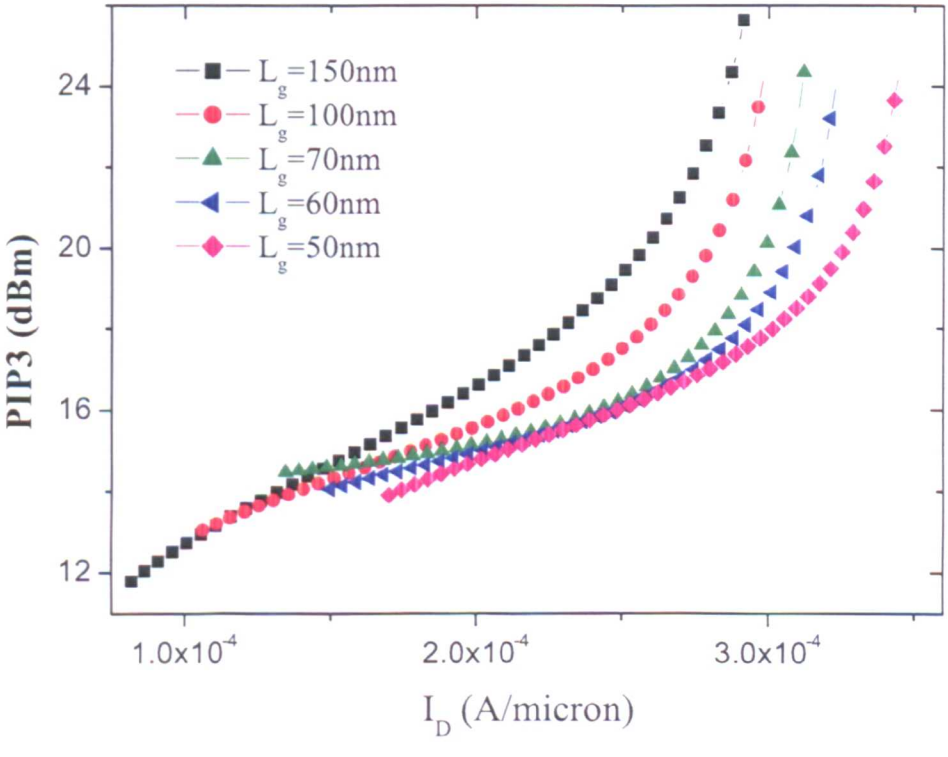


Figure 4.21: The device PIP3 characteristics with different gate length L_g

When the device is scaled laterally, the RF performance is improved as illustrated in Fig. 4.15. Unfortunately, the gate length scaling doesn't help in obtaining a better linearity. For conventional CMOS devices [162], PIP3 starts to decrease when the device is scaled into the deep submicron regime and eventually rises at very small gate length (<100 nm). In this simulation work, the modulation doped device behaves similarly to the traditional MOSFET as illustrated in Fig. 4.21.

The simulations have preserved the vertical structure, mobility, series resistance and saturation velocity during the scaling. If the vertical scaling is also considered, the RF performance may increase even more (see Fig. 4.18) while device linearity will decrease due to a smaller gate-to-channel separation.

4.5.2 Channel Doping Strategies

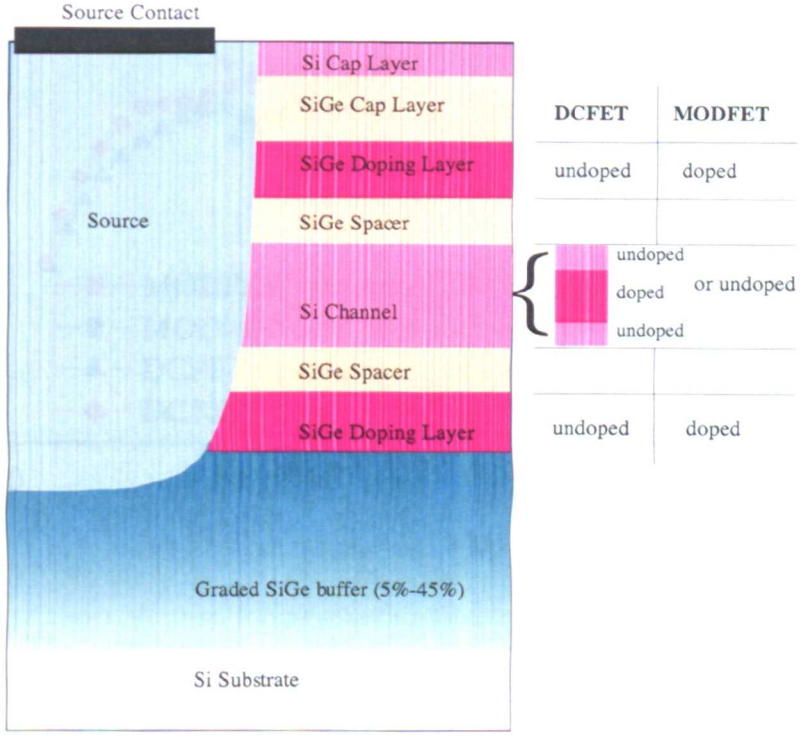


Figure 4.22: Illustrations of studied MODFET and DCFET (half)

The existence of parallel parasitic conduction paths in MODFETs is the major factor affecting linearity. The carriers supplied by the side doping layers may move from the channel to the low mobility parasitic conduction path at high gate voltages,

which narrows the transconductance peak and reduces the linearity. Existing work on III-V HEMTs has suggested that the introduction of channel doping improves linearity [163]. Transferring this idea into Si/SiGe MODFETs, it is also expected to have an increased linearity due to the doped channel. When compared with the side doped devices, the carriers in the doped channel devices stay in the channel for a larger range of gate voltages, which flattens the transconductance characteristics and improves the linearity.

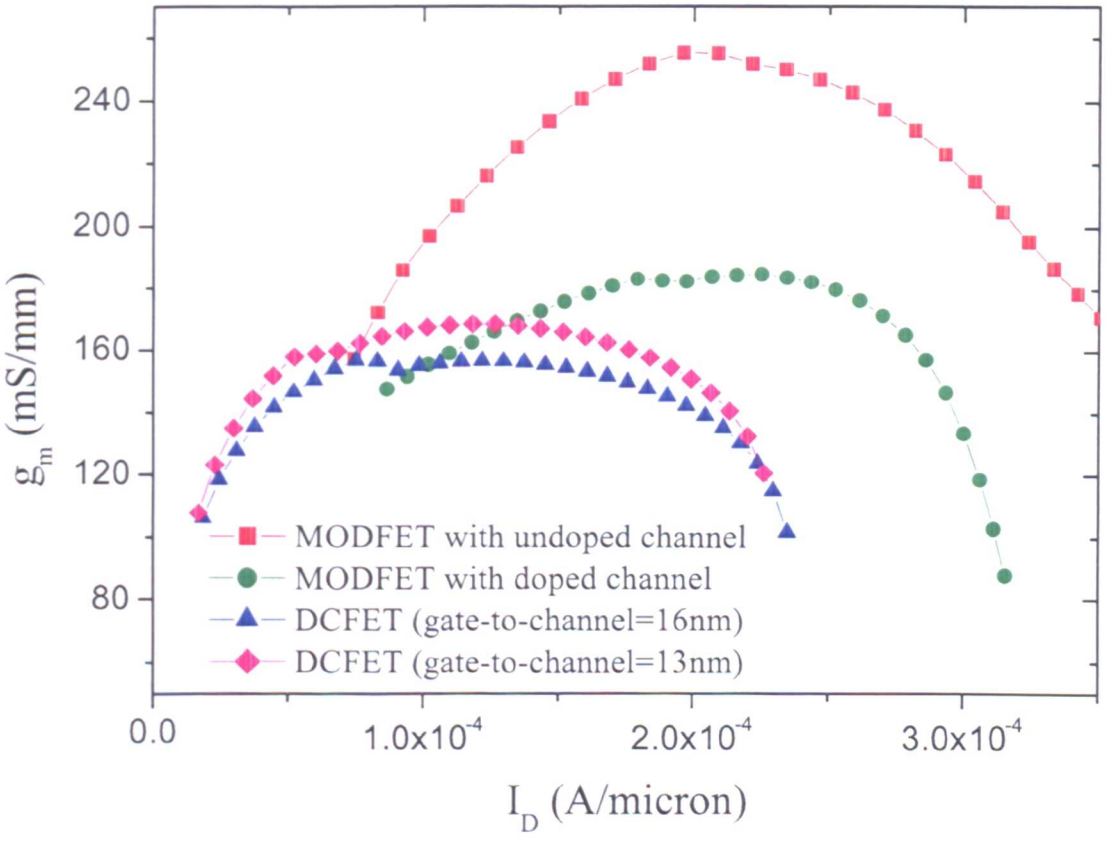


Figure 4.23: The comparison of device transconductance characteristics between structure A (undoped channel MODFET), structure B (doped channel MODFET) and structure C (DCFET)

In this work, three different devices with identical layer structures have been studied (illustrated in Fig. 4.22). Structure A is the original MODFET used in the calibration process (described in section 4.3.2) with the double side SiGe doping layers and an undoped channel. Structure B is the MODFET with a reduced side

doping above the channel compensated by an equivalent amount of channel doping of 1×10^{18} cm^{-3}. Structure C is the doped channel FET (DCFET) without the side doping and with a channel doping of 6×10^{18} cm^{-3}. The doping is limited to the central 5 nm of the channel, as illustrated in Fig. 4.22.

Figs. 4.23 and 4.24 show that the linearity of doped channel devices is distinctly improved compared to that of undoped channel devices while the undoped channel devices have a higher drive current.

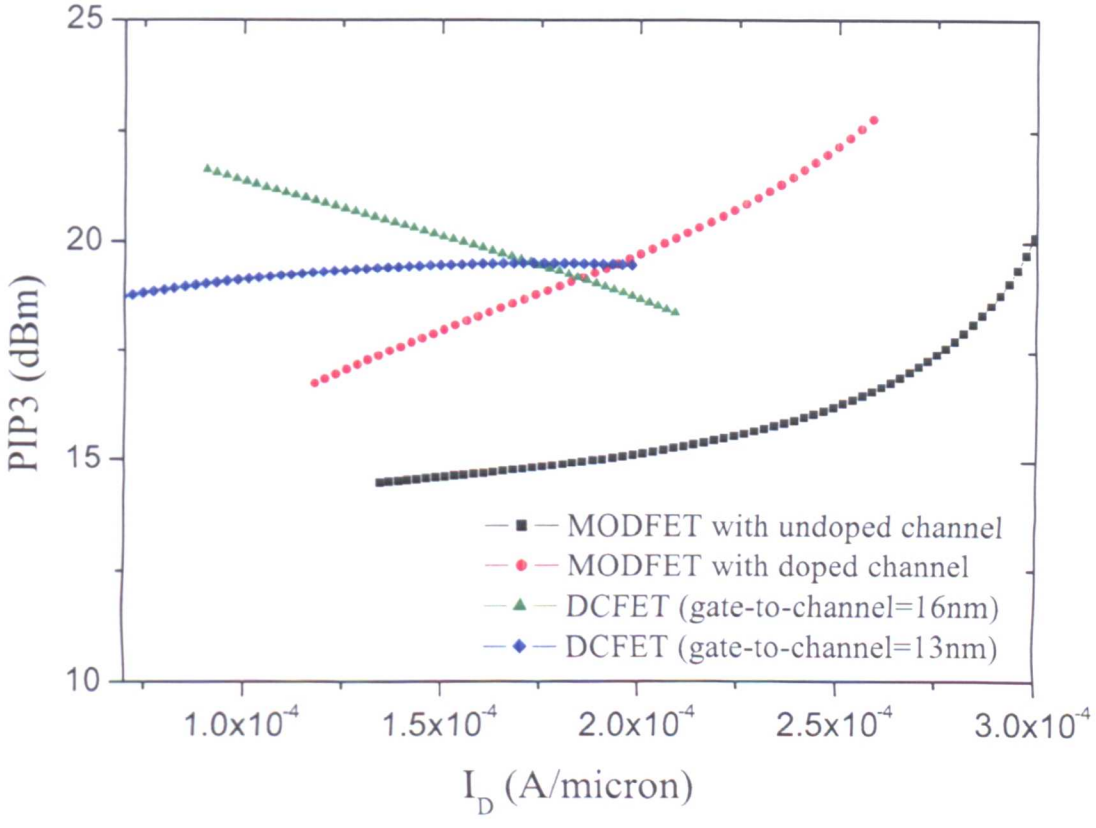


Figure 4.24: The comparison of PIP3 between structure A (undoped channel MODFET), structure B (doped channel MODFET) and structure C (DCFET)

For FET-type amplifiers under small signal operation, distortion related to nonlinearities in the transconductance, g_m, is complemented by distortions associated with nonlinearities in the coupling capacitances, C_{gs}, C_{gd} and C_{ds}. From transient simulations, the voltage dependence of these intrinsic small-signal equivalent circuit parameters (see Fig. 4.2) have been extracted and compared between structure A

(MODFET) and structure C (DCFET), shown in Fig. 4.25. It can be seen from the figure that across a large gate operation range the capacitances of structure C exhibit a greater linearity than that of structure A which contributes to the high linearity of doped channel MODFETs.

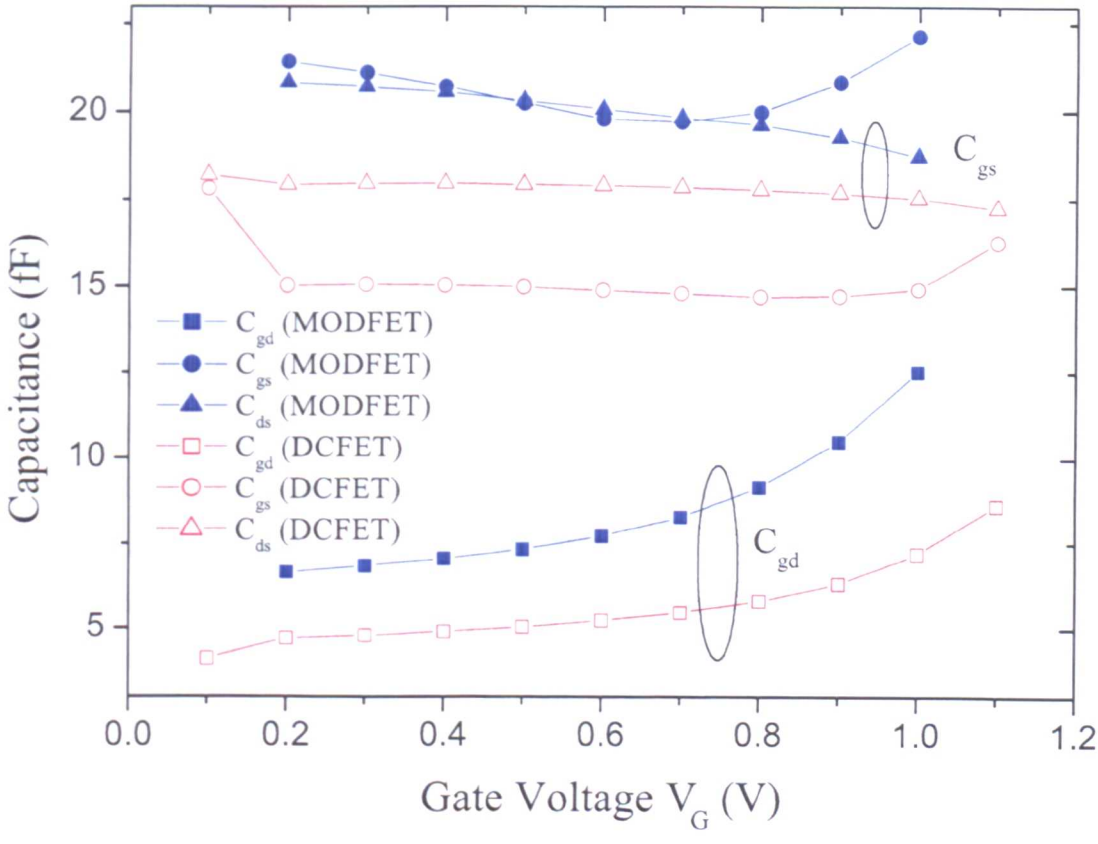


Figure 4.25: The comparison of intrinsic capacitances between structure A (undoped channel MODFET) and structure C (DCFET) at V_D=1.5 V

However, figures 4.23 and 4.24 show that the drive current and the peak transconductance of doped channel MODFETs are much smaller than that of undoped channel devices due to the lower mobility within the channel. Although the doping in the channel provides a high sheet carrier density, the mobility is significantly lower than that of the undoped channel due to strong ionized impurity scattering. By solving Poisson's equation in 1-D , the sheet carrier densities, n_s, in the channel, drain current, I_D, transconductance, g_m, intrinsic capacitance, C_{gs}+C_{gd}, and intrinsic f_T at V_G=0.4 V have been obtained and summarized in Table 4.1. Even with a

Table 4.1: Comparisons of sheet carrier density in the channel n_s ($\times 10^{12}$ cm^{-2}), I_D ($\times 10^{-4}$ A/μm), g_m (mS/mm), intrinsic capacitance (fF) and intrinsic f_T (GHz) between the doped and undoped channel devices

	n_s	I_D	g_m	C_{gs}+C_{gd}	f_T
Structure A	2.18	2.8	231	27.7	83.8
Structure C	1.91	1.4	156	18.9	69.7

much lowered g_m, the doped channel MODFET achieved a relatively high intrinsic cut-off frequency f_T because of the reduced gate capacitance C_{gs}+C_{gd}.

4.6 Summary

Based on the calibration in respect of a 0.25 μm and a 70 nm n-type buried strained Si channel Si/SiGe MODFETs fabricated by Daimler Chrysler, the RF performance and linearity in various Si/SiGe MODFETs architectures have been studied in this chapter. Qualitative analysis and numerical simulations indicate that the RF performance and linearity are sensitive to both lateral and vertical device designs, especially to L_g, L_{ds}, L_{gs}, to the gate-to-channel separation d (which depends on the thickness of each layer) and to the doping concentrations in the SiGe doping supply layer and the channel. The gate-to-channel separation and gate to source/drain distances are found to have a significant but opposite effects on the device performance and linearity. The doping in the supply layers need to be properly adjusted in order to obtain a high sheet carrier density in the channel and to achieve a higher modulation efficiency. The doped channel device exhibits the best linearity but at the expense of the reduced drive current, transconductance and RF performance. The simulations also show that scaling helps to improve RF performance but slightly reduces device linearity. Trade-off designs are necessary for specific RF and/or high linearity applications.

In this chapter, the comprehensive design considerations of Si/SiGe MODFETs originate from the multi-layer structure of their III-V counterparts and the strained Si/SiGe heterostructure. However, the most commonly studied devices of the strained Si on relaxed SiGe heterostructure is its implementation in the structure of conventional MOSFETs for CMOS applications which is addressed in next chapter.

Chapter 5

Scaling Study of Strained Si MOSFETs for CMOS Applications

Sub-100 nm surface channel strained Si/SiGe MOSFETs for CMOS applications have been reported with notable performance enhancements for both n- and p-MOSFETs when compared to conventional Si MOSFETs (discussed in Chapter 2). The performance enhancements are mainly due to strain-induced carrier-mobility enhancement within the strained Si channel and reduced interface roughness scattering in strained Si MOSFETs. However, high strain (>30% Ge content in the buffer) is necessary for strained Si p-type MOS applications in order to obtain a larger hole mobility enhancement (see Fig. 2.4) and match the performance enhancement of n-type strained Si MOSFETs. Moreover, the type II strained Si on relaxed SiGe heterostructure used in the surface channel p-type strained Si MOSFETs (see Fig. 2.1) leads to a parasitic conduction path in the low-mobility relaxed SiGe layer.

In this chapter, extended drift diffusion and hydrodynamic models are used to investigate the performance dependence of p-type strained Si MOSFETs and the scaling properties of the strained Si CMOS technology. The simulations are based on careful calibrations with respect to the sub-100 nm n- and p-type strained Si/SiGe MOSFETs reported by Rim *et al.* [21,31]. The calibrated device structures are then scaled down to 35 nm physical gate lengths to evaluate the potential device and circuit behaviours. The 2D device simulations carried out with MEDICI are com-

plemented with an understanding of the carrier transport of strained Si using Monte Carlo simulations and the parameters associated with the Si/SiGe heterostructure given in Chapter 3.

5.1 Simulation Models and Device Calibrations

5.1.1 Simulation Models

2D drift-diffusion (DDM) and hydrodynamic (HDM) simulations employing MEDICI are used in this chapter to assess the performance of strained Si based devices and circuits. The DDM simulations are based on the models described in Chapter 4, incorporating the concentration dependent model for low-field mobility, the perpendicular field dependent mobility model accounting for surface roughness scattering, and the Caughey-Thomas lateral field dependent model accounting for high field effects. Details of these mobility models have been addressed in Chapter 4. The default parameters in MEDICI associated with the Si/SiGe heterostructure are modified based on the analytical expressions given in Chapter 3.

Similar to the quantum confinement in the channel of Si/SiGe MODFETs discussed in Chapter 4, the surface channel MOSFETs studied in this chapter also feature a 2-dimensional quantum well when high gate voltages are applied. An example of the quantum mechanical effects on the electron distribution in the inversion layer of Si MOSFETs is illustrated in Fig 5.1, which is obtained from Greg Snider's 1D Poisson-Schrödinger solver [28]. It is evident that the quantum confinement in the inversion layer causes a positional shift of the peak electron concentration about 1nm away from the Si/SiO_2 interface. This effect reduces the sheet carrier density in the channel and effectively induces a threshold voltage shift. The quantum correction model used for this study is a bandgap widening approach [143]. The model approximately accounts for both the splitting of energy levels in the conduction band to higher sub-bands and for a displacement of the carrier concentration away from the semiconductor-insulator interface. The bandgap of channel material is therefore widened by an amount $\triangle E_{g,qm}$, as expressed by [143]:

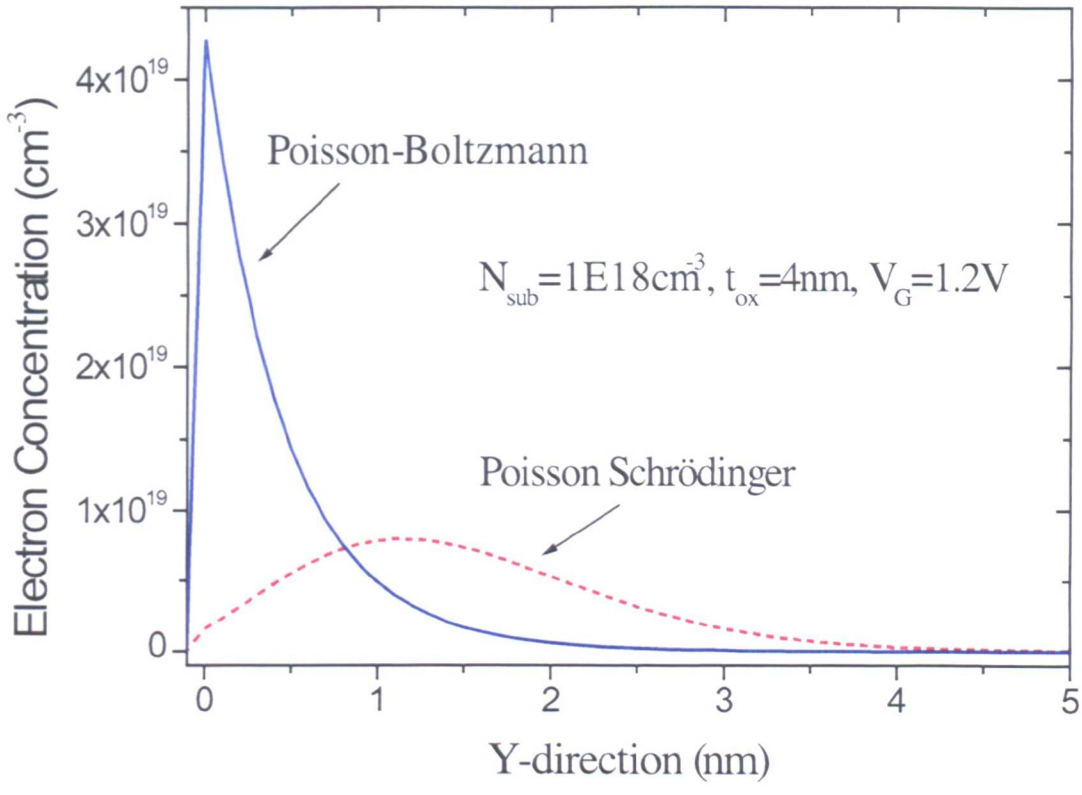


Figure 5.1: Electron distributions solved by 1D Greg-Snider solver [28] based on the solutions of Poisson-Boltzmann and Poisson-Schrödinger equations

$$\triangle E_{g,qm} = \kappa_{n(p)} \cdot \frac{13}{9} \cdot \beta \cdot \left(\frac{\epsilon_{semi}}{4kT}\right)^{1/3} \cdot \left|E_{n(p)}\right|^{2/3} \tag{5.1}$$

where $E_{n(p)}$ is the normal electric field at the semiconductor/insulator interface; β is a fitting parameter and by default set to 4.1×10^{-8} eV·cm; ϵ_{semi} is the permittivity of semiconductor material in the channel; $\kappa_{n(p)}$ is a fitting parameter accounting for the effect of quantized levels above the ground state and by default set to 1.0 for silicon in MEDICI. The value of $\triangle E_{g,qm}$ is then used to calculate the corrected intrinsic carrier concentration at the interface:

$$n_{i,qm} = n_{i,conv} \exp\left(-\frac{\triangle E_{g,qm}}{2kT}\right) \tag{5.2}$$

where $n_{i,conv}$ is the conventional intrinsic carrier concentration at the semiconductor/insulator interface without taking into account quantum mechanical affects.

In addition to the drift-diffusion model discussed in Chapter 4, this chapter also uses hydrodynamic device simulations which approximate the non-equilibrium transport due to local carrier heating at high electric fields. The hydrodynamic model includes the drift-diffusion transport model described by equation (4.1)-(4.3) and the energy balance equation from the second-order moment of the Boltzmann equation. Instead of equations (5.3) and (5.4) used in DDM, the following equations are employed for solving the current densities J_n and J_p [143].

$$\vec{J_n} = q\mu_n \vec{\nabla}(u_n n) + q\mu_n n \vec{E} + qnu_n \frac{\partial \mu_n(u_n)}{\partial u_n} \vec{\nabla} u_n \tag{5.3}$$

$$\vec{J_p} = -q\mu_p \vec{\nabla}(u_p p) + q\mu_p p \vec{E} - qpu_p \frac{\partial \mu_p(u_p)}{\partial u_p} \vec{\nabla} u_p \tag{5.4}$$

where u_n and u_p represent the electron and hole thermal voltages kT_n/q and kT_p/q respectively (T_n and T_p are the electron and hole temperatures). Neglecting carrier generation and recombination, the energy balance equations for electrons and holes may be expressed by [143]:

$$\vec{\nabla} \cdot \vec{S_n} = \frac{1}{q}\vec{J_n} \cdot \vec{E} - \frac{3}{2}\left[n\frac{u_n - u_0}{\tau_n} + \frac{\partial(nu_n)}{\partial t}\right] \tag{5.5}$$

$$\vec{\nabla} \cdot \vec{S_p} = \frac{1}{q}\vec{J_p} \cdot \vec{E} - \frac{3}{2}\left[p\frac{u_p - u_0}{\tau_p} + \frac{\partial(pu_p)}{\partial t}\right] \tag{5.6}$$

where u_0 represents the lattice thermal voltage kT_0/q (T_0 is the lattice temperature); τ_n and τ_p denote the electron and hole energy relaxation times and may be obtained from Monte Carlo simulations; S_n and S_p represent the electron and hole energy flow densities and are given by [143]:

$$\vec{S_n} = -\frac{5}{2}u_n\left[\frac{\vec{J_n}}{q} + \mu_n n \vec{\nabla} u_n\right] \tag{5.7}$$

$$\vec{S_p} = \frac{5}{2}u_p\left[\frac{\vec{J_p}}{q} - \mu_p p \vec{\nabla} u_p\right] \tag{5.8}$$

The determined carrier temperature distributions make it possible to more accurately model the real carrier transport. In this case, the Caughey-Thomas field dependent mobility model, given by equation (5.9) for the DDM simulations, depends on the carrier and lattice temperatures and the carrier energy relaxation times and becomes [143]:

$$\mu_{n(p)}(u_{n(p)}) = \frac{\mu_{S,n(p)}}{\left[1 + \left[\alpha_{n(p)}(u_{n(p)} - u_0)\right]^{\beta}\right]^{1/\beta}} \tag{5.9}$$

$$\alpha_{n(p)} = \frac{3\mu_{S,n(p)}}{2v_{sat}^2 \tau_{n(p)}} \tag{5.10}$$

5.1.2 Device Calibrations of sub-100nm Strained Si MOSFETs

Using the DDM model, MEDICI simulations are carried out to perform device calibrations in respect of sub-100 nm conventional Si and strained Si n-type and p-type MOSFETs. The calibrated n-type devices have an 80 nm gate length (L_g) and a 2.2 nm thick gate oxide (t_{ox}) for both Si and strained Si MOSFETs and a relaxed SiGe *virtual substrate* with 15% Ge content is used in the strained Si device [31]. Both calibrated p-type Si and strained Si MOSFETs have a 90 nm gate length (L_g) and a 2.8 nm thick gate oxide (t_{ox}), and the strained Si device uses a relaxed SiGe *virtual substrate* with 28% Ge content [21]. The device structure of the n-type strained Si MOSFET is illustrated in Fig. 5.2. The p-type strained Si MOSFET has the same schematic as the n-type strained Si device. The vertical band diagram of the device is also plotted in Fig. 5.2. The band offsets at the type-II Si/SiGe hetero-interface enable the electron confinement in the strained Si layer, but lead to a parasitic conduction path for holes in the low-mobility SiGe layer. The Si devices are assumed to have the same physical dimensions and identical source/drain doping profiles as their strained Si counterparts. Channel doping profiles between the Si and strained Si MOSFETs are slightly different due to the different diffusion rates of dopants in Si and in SiGe [74, 75], although the same process conditions are assumed for both devices.

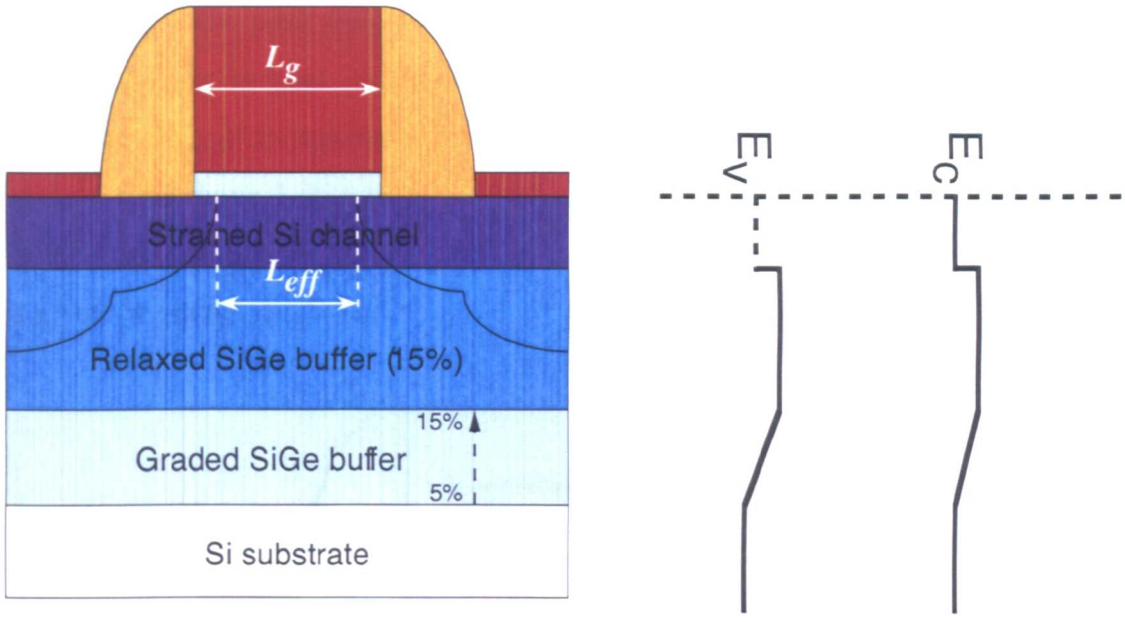


Figure 5.2: Schematic of strained Si MOSFET and band diagram along the middle of the device

The calibrations against the experimental data of these devices are based on the methodology described in Chapter 4. The calibration uses a single retrograde well doping and a n-type heavily doped poly-Si gate for both Si and strained Si MOSFETs. Gaussian distribution based analytical profiles are used to reproduce the source, drain and channel doping distributions. An illustration of the doping profiles in the devices is shown in Fig. 5.3. Assuming a symmetric doping distribution in the lateral direction (x-direction) and setting up a $x-y$ coordinate in the device (shown in Fig. 5.3), the doping concentration at a mesh point (x, y) may be expressed by a sum of several 2D analytical profiles:

$$N_{total}(x,y) = N_{D1} f_{x1}(x) f_{y1}(y) + N_{D2} f_{x2}(x) f_{y2}(y) + N_{sub} + N_A f_{y3}(y) \tag{5.11}$$

$$f(z) = \begin{cases} \exp\left[-\left(\frac{z-z_{min}}{z_{char}}\right)^2\right] & z < z_{min} \\ 1 & z_{min} < z < z_{max} \\ \exp\left[-\left(\frac{z-z_{max}}{z_{char}}\right)^2\right] & z > z_{max} \end{cases} \tag{5.12}$$

where N_{D1} is the peak concentration of the n-type doping in the source (or drain) contact; N_{D2} is the peak concentration of the n-type doping in the source (or drain) extension; N_{sub} is the uniform substrate p-type doping; N_A is the peak concentration of the p-type channel doping; f_{xi} and f_{yi} are the horizontal (x-direction) and vertical (y-direction) doping variation functions. The general expression of the function $f(z)$ is given by equation (5.12).

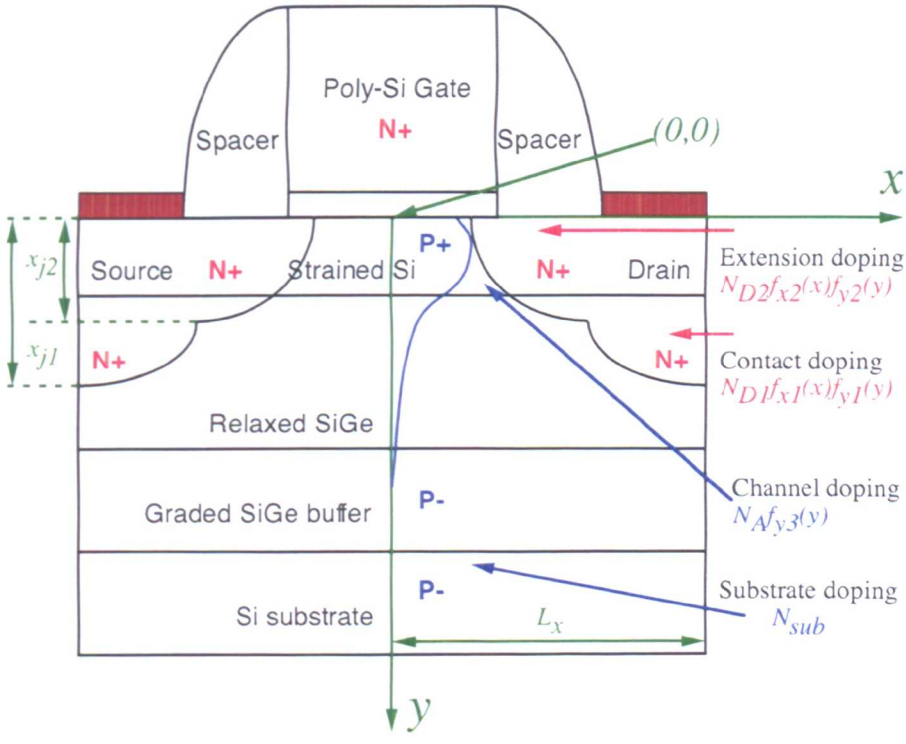


Figure 5.3: Illustration of the doping profiles within the strained Si MOSFET

A 2D doping distribution may be obtained by expression (5.11) through a combination of N_A, N_D, x_{char}, y_{char}, x_{min}, x_{max}, y_{min} and y_{max}. The source/drain doping parameters are obtained by matching the expression with the available doping profiles [71] and the channel doping parameters are calculated by calibrating with respect to the subthreshold slope and the threshold voltage of the experimental data. The calibrated doping distribution parameters for both Si and strained Si MOSFETs are listed in Table 5.1.

Table 5.1: Summary of calibrated doping profiles in the 80 nm Si and strained Si n-MOSFETs

	Units	Si	Strained Si
N_{sub}	cm^{-3}	1×10^{15}	1×10^{15}
Channel doping: $N_A(cm^{-3})$	cm^{-3}	3×10^{18}	2×10^{18}
Channel doping: y_{char}	nm	60	50
Channel doping: y_{min}/y_{max}	nm/nm	30/30	40/40
S/D contact doping: N_{D1}	cm^{-3}	2×10^{20}	2×10^{20}
S/D contact doping: x_{char}/y_{char}	nm/nm	17/25	17/25
S/D contact doping: x_{min}/x_{max}	nm/nm	$100/L_x$	$100/L_x$
S/D contact doping: y_{min}/y_{max}	nm/nm	2/2	2/2
S/D extension doping: N_{D2}	cm^{-3}	2.5×10^{20}	2.5×10^{20}
S/D extension doping: x_{char}/y_{char}	nm/nm	9/14	9/14
S/D extension doping: x_{min}/x_{max}	nm/nm	$47/L_x$	$47/L_x$
S/D extension doping: y_{min}/y_{max}	nm/nm	2/2	2/2

L_x is half of the total length in the x-direction of the simulated devices

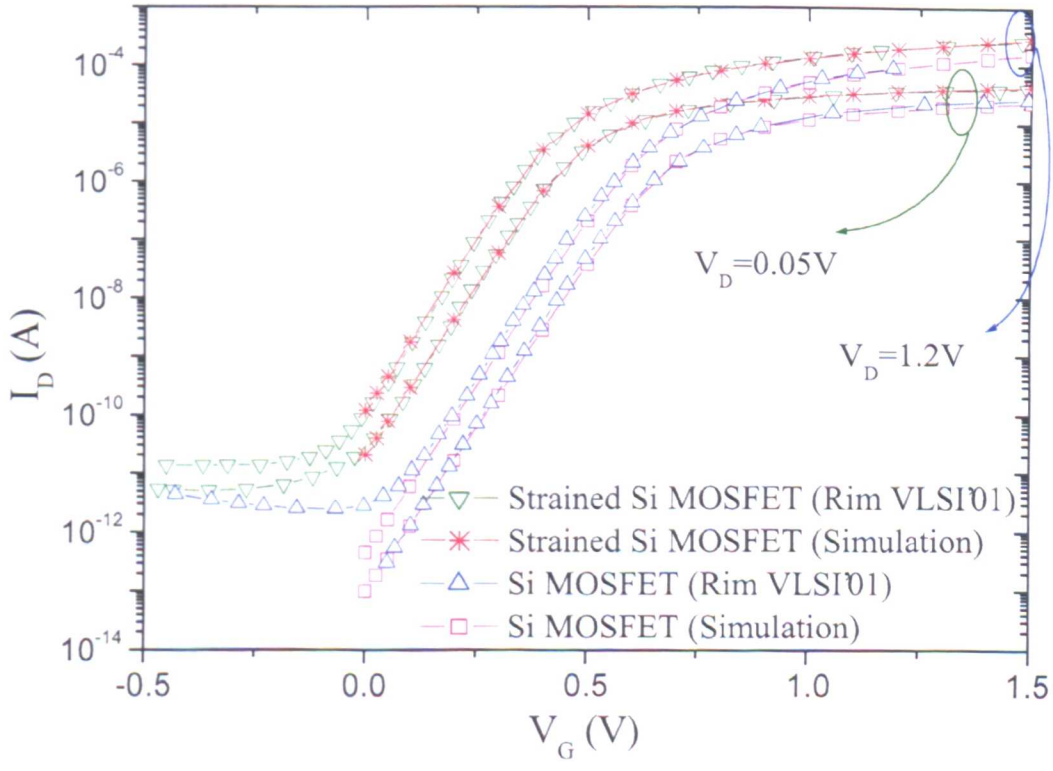


Figure 5.4: Calibrated I_D-V_G characteristics of the 80 nm gate length Si and strained Si n-MOSFETs

Fig. 5.4 illustrates the calibrated I_D-V_G characteristics of the 80 nm gate length Si and strained Si n-MOSFETs. The simulated I_D-V_G characteristics agree with the experimental data across whole device operation range for both Si and strained Si devices. A 200 mV threshold voltage difference is observed between the two MOSFETs which is due to the change of band structure of strained Si and the suppressed Boron diffusion of the channel doping within the strained Si MOSFET [31, 75]. The successful calibrations of both Si and strained Si MOSFETs indicate that the reproduction of the device structures and the doping profiles provides reliable information for later simulation studies.

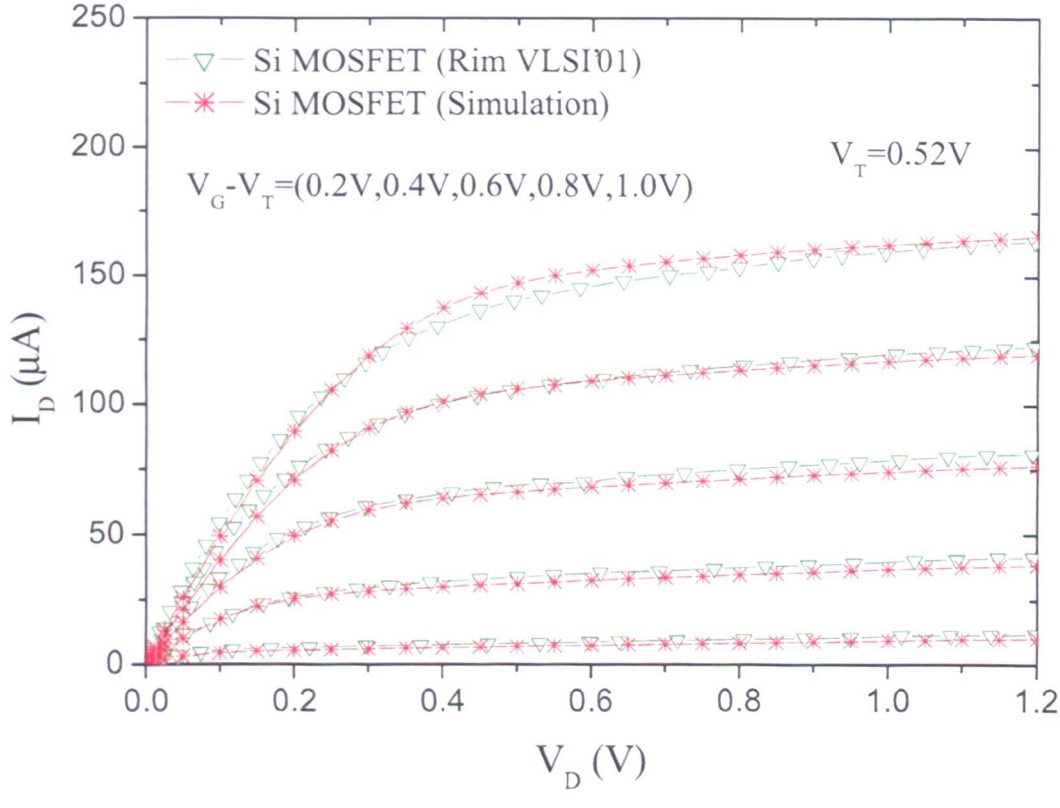


Figure 5.5: Calibrated I_D-V_D characteristics of the 80 nm gate length Si n-MOSFET

If we define an effective gate length (L_{eff}) as the distance between the points where the source/drain doping falls to 2×10^{19} cm^{-3} [164], the effective gate length of the 80 nm gate length Si and strained Si MOSFETs is about 67 nm. Using I_D-V_G characteristics, we may also define a threshold voltage V_T as the gate voltage at which

the drain current equals $L_{eff} \times 10^{-7}$A/μm at low drain voltage (50 mV) [165]. The extracted threshold voltage is 0.52 V in the Si MOSFET and 0.32 V in the strained Si MOSFET. The extracted subthreshold slope, which is defined as $S = \left(\frac{d(log_{10} I_D}{dV_G}\right)^{-1}$ [165], is 86.7 mV/decade for the Si MOSFET and 88.3 mV/decade for the strained Si MOSFET. These values indicate well controlled short channel effects in both devices. The drive current enhancement of the strained Si MOSFET, compared to the Si device, is around 35% [31] at the same gate overdrive V_G-V_T=1.0 V. Figs. 5.5 and 5.6 plot the calibrated I_D-V_D characteristics of the Si and strained Si MOSFETs. Consistent with the I_D-V_G characteristics, the simulated I_D-V_D device characteristics agree with the experimental data throughout the whole range of device operation giving even more confidence to the calibration.

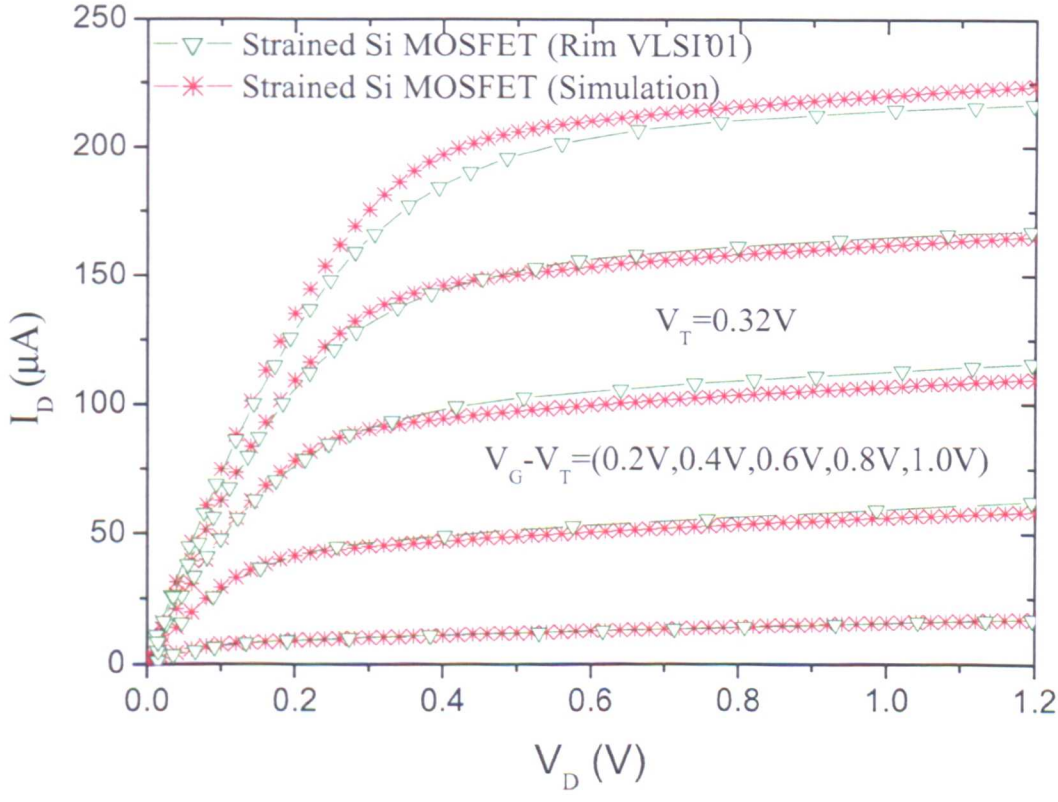


Figure 5.6: Calibrated I_D-V_D characteristics of the 80 nm gate length strained Si n-MOSFET

The calibrations of the 90 nm Si and strained Si p-type MOSFETs follow the same methodology as that of the n-type MOSFETs. The devices have the same effective

gate length of about 67 nm. Fig. 5.7 illustrates the simulated I_D-V_G characteristics of both devices, compared to experimental data from [21]. Slight discrepancies in the subthreshold region of both devices shown in Fig. 5.7 may be due to the lack of reliable information concerning the doping profiles. The doping distributions used for the initial guess of the calibration are based on the information of the calibrated n-type devices. The extracted subthreshold slope S is approximately 92 mV/decade for both the Si and strained Si MOSFETs. The extracted threshold voltage V_T is approximately -0.35 V for the Si MOSFET and -0.3 V for the strained Si MOSFETs. The small difference in the threshold voltages between the p-type Si and strained MOSFETs, compared to that of n-type devices, benefits from the smaller change of the band energy level in the valence band of strained Si and the small difference of the channel doping distributions between the Si and strained Si devices. The strained Si MOSFET delivers about 10% drive current enhancement over the Si MOSFET. Table 5.2 summarizes the calibrated doping distributions parameters for the 90 nm Si and strained Si p-MOSFETs.

Table 5.2: Summary of calibrated doping profiles for the 90 nm Si and strained Si p-MOSFETs

	Units	Si	Strained Si
N_{sub}	cm^{-3}	1×10^{15}	1×10^{15}
Channel doping: $N_A(cm^{-3})$	cm^{-3}	2×10^{18}	2.15×10^{18}
Channel doping: y_{char}	nm	80	80
Channel doping: y_{min}/y_{max}	nm/nm	35/35	35/35
S/D contact doping: N_D	cm^{-3}	2×10^{20}	2×10^{20}
S/D contact doping: x_{char}/y_{char}	nm/nm	21/30	21/30
S/D contact doping: x_{min}/x_{max}	nm/nm	$110/L_x$	$110/L_x$
S/D contact doping: y_{min}/y_{max}	nm/nm	2/2	2/2
S/D extension doping: N_D	cm^{-3}	2.5×10^{20}	2.5×10^{20}
S/D extension doping: x_{char}/y_{char}	nm/nm	10/14	10/14
S/D extension doping: x_{min}/x_{max}	nm/nm	$45/L_x$	$45/L_x$
S/D extension doping: y_{min}/y_{max}	nm/nm	2/2	2/2

L_x is half of the total length in the x-direction of the simulated devices

Based on the calibrations in respect of the 80 nm n-type Si and strained Si MOSFETs and the 90 nm p-type Si and strained Si MOSFETs, more extensive

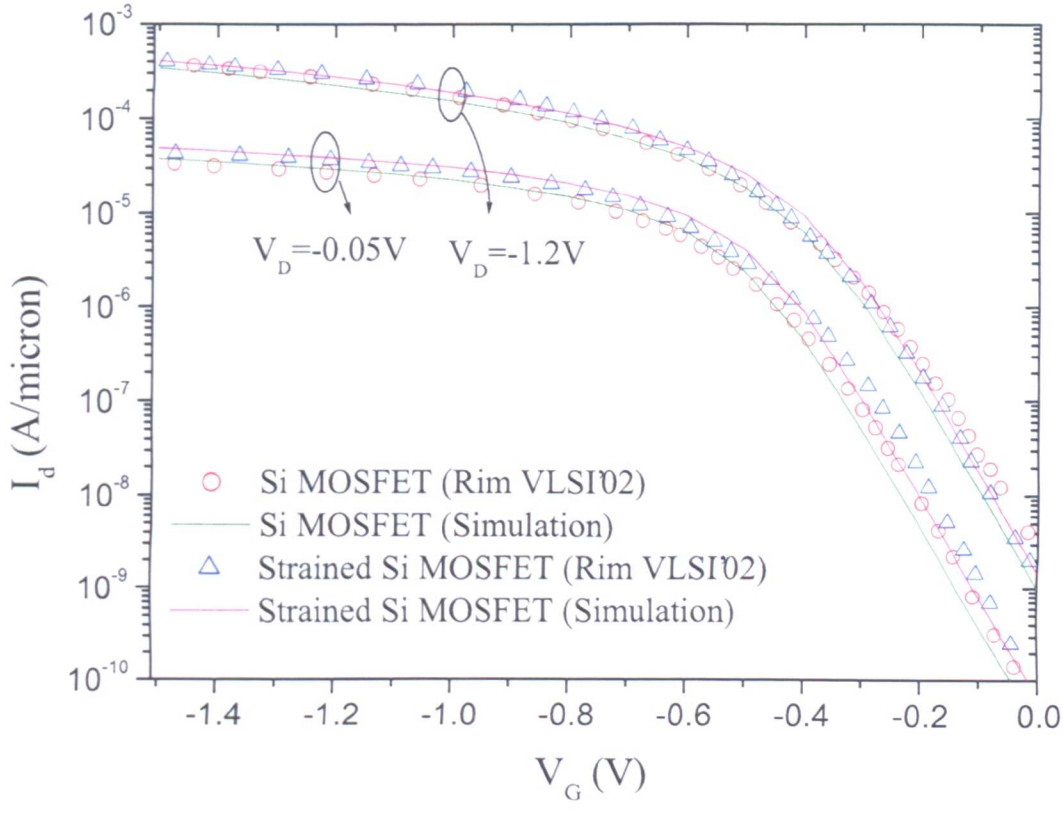


Figure 5.7: Calibrated I_D-V_G characteristics of the 90 nm gate length Si and strained Si p-MOSFETs

work may now be carried out to assess the behaviour of sub-100 nm strained Si *MOSFETs for CMOS applications.*

5.2 Impact of Parasitic Channel in p-type Strained Si MOSFETs

As discussed in Chapter 2, to balance the performance between the n-type and p-type devices for CMOS applications, a high Ge content may be required for p-type strained Si MOSFETs. However, one problem of the strained Si/SiGe heterostructure for p-type MOSFET applications is that the valence band offset tends to confine holes in the low mobility SiGe layer (see Fig. 2.1), affecting the device characteristics and increasing the leakage current. Fig. 5.8 compares the hole distribution and valence band profile for a range of uniform substrate doping levels obtained from

the Greg-Snider 1-D Poisson-Schrödinger solver [28]. It is clear that the increase in the substrate doping suppresses the hole population within the $Si_{0.65}Ge_{0.35}$ layer. When N_A exceeds 10^{17} cm^{-3}, the hole distribution in the parasitic channel (SiGe layer) becomes negligible.

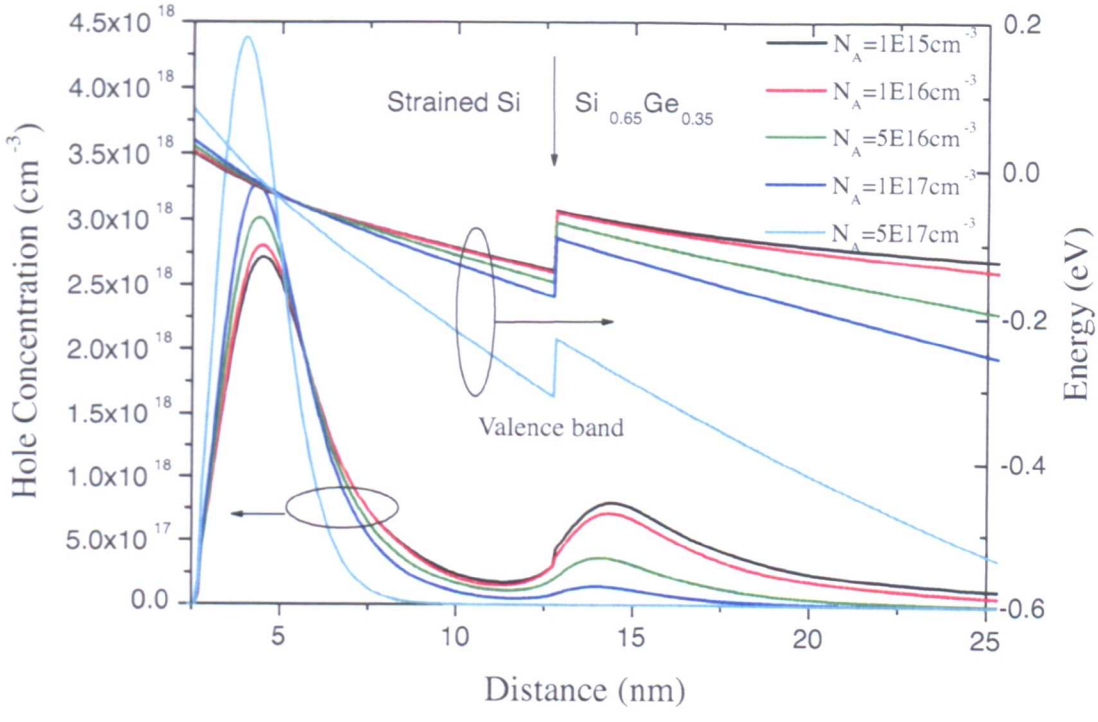


Figure 5.8: Hole distributions and valence band profile in the channel for a range of uniform substrate dopings obtained using 1-D Poisson-Schrödinger solutions (sheet densities in the channel n_s=1×10^{12} cm^{-2})

The hole density may also be reduced by lowering the Ge concentration of the SiGe buffer. However, it is crucial to maintain a high Ge concentration in the buffer as the performance improvement of strained Si MOSFET is rapidly reduced at high V_G. For example, in the calibrated 90 nm p-type strained Si MOSFET, the mobility enhancement of strained Si on $Si_{0.72}Ge_{0.28}$ drops from a maximum of 45% and disappears at the vertical effective field approaching to 1 MV/cm [21]. Therefore, a high Ge concentration in the SiGe buffer is required to achieve mobility enhancement at high electric fields.

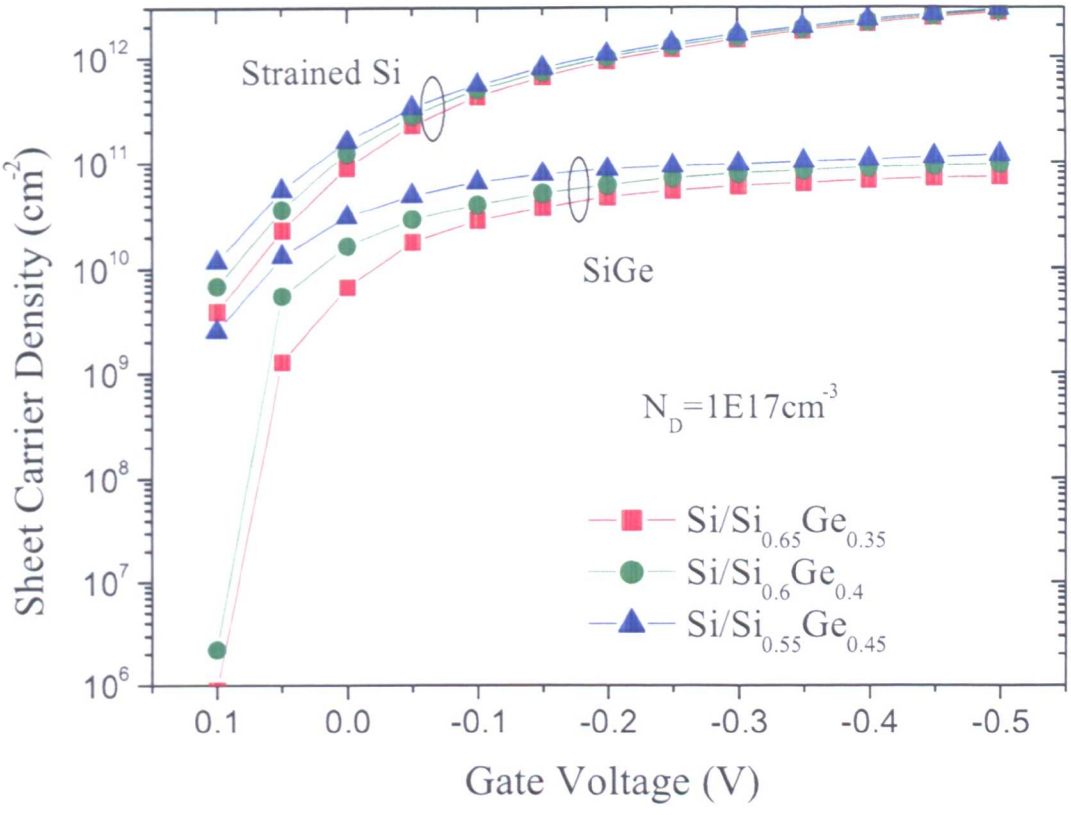


Figure 5.9: Sheet carrier densities in the channel and the SiGe buffer layer with different Ge concentrations as a function of gate voltage obtained using a 1-D Poisson-Schrödinger solver (assuming a uniform substrate doping N_D=1×10^{17} cm^{-3})

Fig. 5.9 shows the sheet carrier densities in the strained Si and SiGe layers for three different Ge concentrations in the buffer. At high gate voltages, the carrier density in the SiGe buffer becomes negligible compared to that in the channel. At low gate bias, although increasing the Ge content increases hole concentration at the buffer-channel interface, it remains smaller than the hole concentration in the strained Si channel due to the high channel doping and does not compromise the subthreshold leakage. Using HDM device simulations, Fig. 5.10 illustrates the simulated I_D-V_G characteristics of devices with different Ge concentrations. The hole energy relaxation times, used in the HDM simulations and obtained from Monte Carlo, are 0.2 ps, 0.69 ps, 0.81 ps and 0.9 ps for strained Si on relaxed SiGe with Ge content of 0% (bulk Si), 28%, 35% and 40%, respectively.

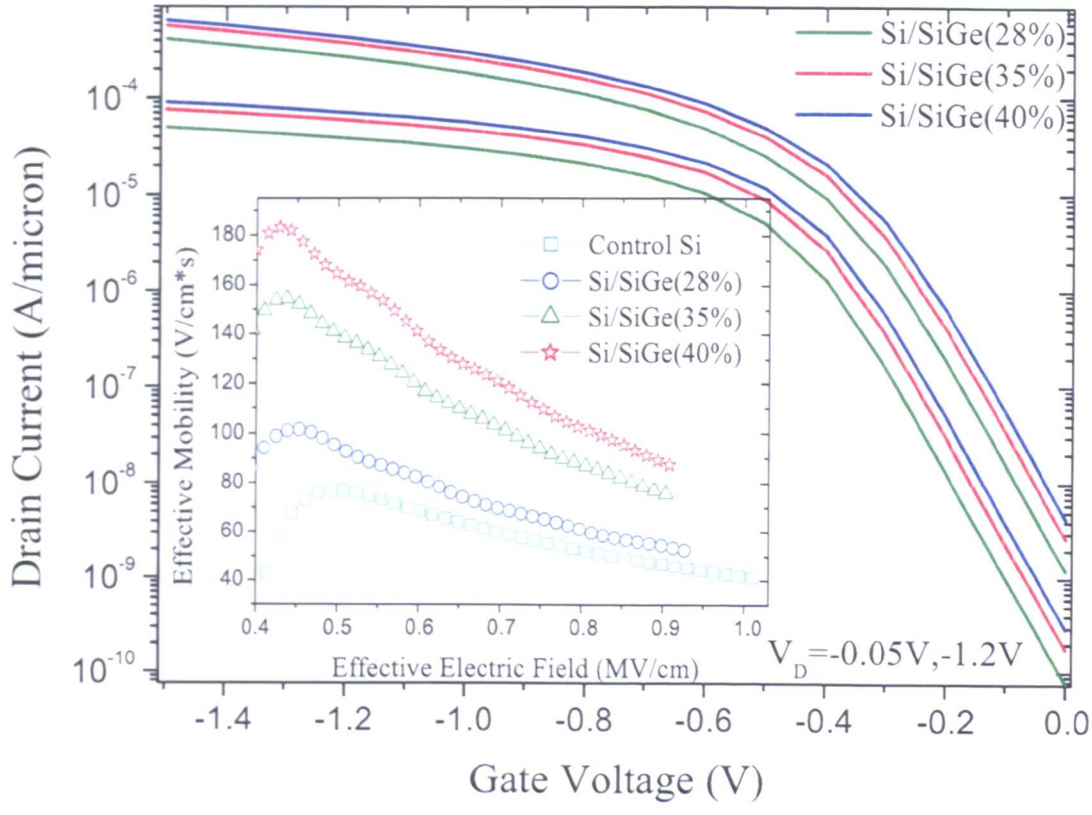


Figure 5.10: Simulated I_D-V_G characteristics of 90 nm strained Si MOSFETs with different Ge contents in the buffer; the inset is the extracted effective mobility from the simulations

In the simulations, the mobility models have been calibrated to obtain the same effective mobility enhancement factors as available experimental data [72] which is around 2 for 35% Ge content and around 2.25 for 40% Ge content at the effective electric field E_{eff} =0.5 MV/cm, and around 2 for 40% Ge content at E_{eff}=0.6 MV/cm. The inset of Fig. 5.10 shows the extracted effective mobilities from the MEDICI simulations. Extractions of the effective mobility μ_{eff} and the the effective electric field E_{eff} are based on the theory used by Currie *et al.* [110]. At small V_D (50 mV in this case), E_{eff} is calculated from $E_{eff} = \frac{E_{ox}\epsilon_{ox}-(1-\eta)Q_{inv}}{\epsilon_S}$ and μ_{eff} is extracted from $\mu_{eff} = \frac{I_D L}{W C_{ox}(V_G-V_T)V_D}$. In these expressions, E_{ox} is assumed to be equal to V_G/t_{ox}; Q_{inv} is the inversion layer charge taken as $C_{ox}(V_G - V_T)$; ϵ_{ox} and ϵ_S are the dielectric constants of the oxide and the semiconductor (strained Si in this case); η is a fitting parameter and typically set to 1/3 for holes in order to attain the universal mobility behaviour; L and W are the device effective gate length and

the device width respectively; C_{ox} is the effective gate oxide capacitance and taken as ϵ_{ox}/t_{inv}; t_{inv} is the equivalent oxide thickness which equals to the sum of the physical gate oxide thickness (t_{ox}) and the equivalent oxide thickness contributed by the inversion layer.

In Fig. 5.10, the subthreshold slope is only weakly affected by the Ge content of the SiGe buffer, indicating that the high channel doping ($>10^{17}$ cm^{-3}) suppresses the effect of the parasitic channel. The threshold voltage (V_T) shift in the same figure comes from the change in the valence band offset. The increase in saturation current with increasing Ge content at V_G-V_T−1 V (which shows a 10% increase for a Ge content of 28%; a ~30% increase for Ge content of 35% and a ~50% increase for a Ge content of 40%) is due to the higher sheet carrier density (see Fig. 5.9) and enhanced hole transport. However, it has been speculated that increased misfit dislocation density may affect the mobility enhancement with Ge content above 40% [110].

It is therefore concluded that for sub-100 nm p-type MOSFETs, a buffer Ge content between 30% and 40% is suitable in conjunction with high doping ($>10^{17}$ cm^{-3}) in the channel region in order to suppress parasitic conduction.

5.3 Device Behaviour of Scaled Strained Si MOSFETs

5.3.1 Scaling of n-type and p-type strained Si MOSFETs

After successful calibrations of the sub-100 nm n-type and p-type strained Si MOSFETs, it is possible to investigate the scalability of these devices with a single retrograde channel doping. The calibrated n-type and p-type devices both have effective gate lengths of around 67 nm, which corresponds to the in-production 130 nm technology node in the ITRS roadmap [1]. Using the drift-diffusion device simulations, these devices are scaled down to 45 nm, 35 nm and 25 nm effective gate lengths, corresponding to the 100 nm (which appeared in 2003), 80 nm (expected in 2005) and 65 nm (expected in 2007) technology nodes [1] respectively. The scaling study

is based on the generalized scaling rule [165], which allows different scaling factors for power supply voltage and device dimensions. A simple description of the scaling rule is given below [165].

Assuming that the electric field intensity changes by a factor of α and the device physical dimensions scale down by a factor of κ (κ>1), the potential will change by a factor of α/κ. Therefore, both the vertical and the lateral electric fields change by the same multiplication factor so that the shape of the electric field pattern is preserved. Moreover, the doping concentrations must be scaled up by a factor of $\alpha\kappa$ to control the depletion region depth and avoid short-channel effects due to the higher electric field. This assures sufficient control of 2-D effects when scaling to smaller dimensions.

As a result, the circuit delay scales down by a factor between κ and $\alpha\kappa$, depending on the degree of velocity saturation. The scaling factor of the power-delay product per circuit is α^2/κ^3. However, the main problems of the generalized scaling are the increase of electric field by a factor of α, which causes reliability concerns, and the increase of power density by a factor of α^2 to α^3, which leads to problems in chip packaging.

In this study, the channel doping concentrations of the calibrated 80 nm n-type and 90 nm p-type strained Si MOSFETs have been changed in order to match the threshold voltages of their conventional Si counterparts and enable fair comparisons of devices and circuits. During the scaling processes, the same geometry and source/drain doping profiles are assumed for both conventional and strained Si MOSFETs. The scalings of these threshold voltage matched (V_T-matched) n-type and p-type conventional Si and strained Si MOSFETs are summarized in Tables 5.3 and 5.4.

Table 5.3: Scaling summary of the n-type V_T-matched conventional Si and strained Si MOSFETs

Gate length L_g (nm)		80	65	45	35
Effective gate length L_{eff} (nm)		67	45	35	21
Supply voltage V_{DD} (V)		1.2	1.1	1.0	0.9
Gate oxide thickness t_{ox} (nm)		2.2	1.8	1.4	1.1
SDE junction depth x_{j2} (nm)		31	25	20	17
SDE doping abruptness x_{js} (nm/decade)		15	12	9.5	8
Spacer width (nm)		60	40	32	25
Threshold voltage V_T (V) at V_D–50mV	Si	0.56	0.49	0.42	0.35
	Strained Si	0.56	0.49	0.42	0.35
Threshold voltage roll off ΔV_T (mV)	Si	58	70	90	110
	Strained Si	50	58	75	95
On Current I_{on} (A/m) at V_D–V_G–V_{DD}	Si	322	380	450	525
	Strained Si	400	480	550	647
Off Current I_{off} (pA/μm) at V_D–V_{DD}	Si	1.39	12	203	3300
	Strained Si	1.56	13	147	2570
Subthreshold slope S (mV/decade) at V_D–V_{DD}	Si	85.3	85.2	87.6	92.2
	Strained Si	86.7	86	87.4	89.8

Table 5.4: Scaling summary of the p-type V_T-matched conventional Si and strained Si MOSFETs

Gate length L_g (nm)		90	65	45	35
Effective gate length L_{eff} (nm)		67	45	35	22
Supply voltage V_{DD} (V)		-1.2	-1.1	-1.0	-0.9
Gate oxide thickness t_{ox} (nm)		2.8	2	1.4	1.1
SDE junction depth x_{j2} (nm)		33	25	20	16
SDE doping abruptness x_{js} (nm/decade)		16	12	9.5	8
Spacer width (nm)		65	45	35	30
Threshold voltage V_T (V) at V_D–-50mV	Si	-0.55	-0.47	-0.41	-0.34
	Strained Si	-0.55	-0.47	-0.41	-0.34
Threshold voltage roll off ΔV_T (mV)	Si	130	120	110	110
	Strained Si	120	110	105	100
On Current I_{on} (A/m) at V_D–V_G–V_{DD}	Si	149	168	200	240
	Strained Si	160	190	225	262
Off Current I_{off} (pA/μm) at V_D–V_{DD}	Si	44	100	519	6120
	Strained Si	38	93	490	4300
Subthreshold slope S (mV/decade) at V_D–V_{DD}	Si	91.9	88.9	88.9	94.8
	Strained Si	91.8	88.4	87.3	90.8

The device structure parameters listed in Tables 5.3 and 5.4 may be referred to Fig. 5.3. SDE in the tables stands for the source/drain extension shown in Fig. 5.3. SDE doping abruptness, x_{js}, is defined as the vertical abruptness of the SDE doping.

The scaling processes decrease all the lateral and vertical device dimensions, while increasing the doping concentrations and leading to abrupt junctions. Although the abrupt (with x_{js}–8 nm/decade) and shallow (with x_{j2}–16 nm) SDE junctions in the scaled 35nm devices is achievable in modern CMOS technology [166], they may degrade the drive current due to increased external resistance and poor SDE to gate coupling [167]. Therefore, in reality, such device structure with single retrograde channel doping is not recommended for ultra small devices [1].

As summarized in Tables 5.3 and 5.4, the conventional and strained Si devices show similar behaviour in the subthreshold regime during the scaling process. The devices behave reasonably well across all gate lengths studied, with subthreshold slopes of 87-100 mV/decade and threshold voltage roll off of 50-130 mV. The off-state currents of the scaled devices are lower than the requirements set out in the ITRS roadmap [1]. This indicates that such device structure with single retrograde channel doping may be scaled down to 35 nm gate length with well controlled short channel effects. The drive current enhancement factors, *i.e.*, the ratio between the on currents of the strained Si MOSFETs and the conventional Si MOSFETs, remain the same during the scaling for both n-type and p-type devices. This is because although the DDM model used here is able to predict the subthreshold device behaviours, it doesn't account for the increasing non-equilibrium transport when scaling the device down. Therefore, extended device simulations with more appropriate models, such as the hydrodynamic model and Monte Carlo, are required to fully estimate the device performance of the scaled devices.

5.3.2 Hydrodynamic device simulations of scaled p-type strained Si MOSFETs

In contrast to the successfully demonstration of n-type strained Si MOSFET with gate length down to 25 nm regime [23], little information is available on the performance of scaled p-type strained Si MOSFETs with gate length less than 50 nm. From Table 5.4, the predicted drive current enhancement of the scaled 35 nm p-type strained Si MOSFET over the Si MOSFET is only around 10%, which is similar to the enhancement observed in the original 90nm gate length devices. However, when

considering non-equilibrium transport which the DDM model cannot account for, a greater performance enhancement in the scaled strained Si MOSFETs is to be expected. Here, based on the scaling scenario described in Table 5.4, hydrodynamic simulations are carried out to assess the scaling properties of sub-100 nm strained Si p-channel MOSFETs. The hole energy relaxation times, used in the HDM simulations and obtained from Monte Carlo, are 0.2 ps and 0.69 ps for bulk Si and strained Si on relaxed SiGe with Ge content of 28% respectively.

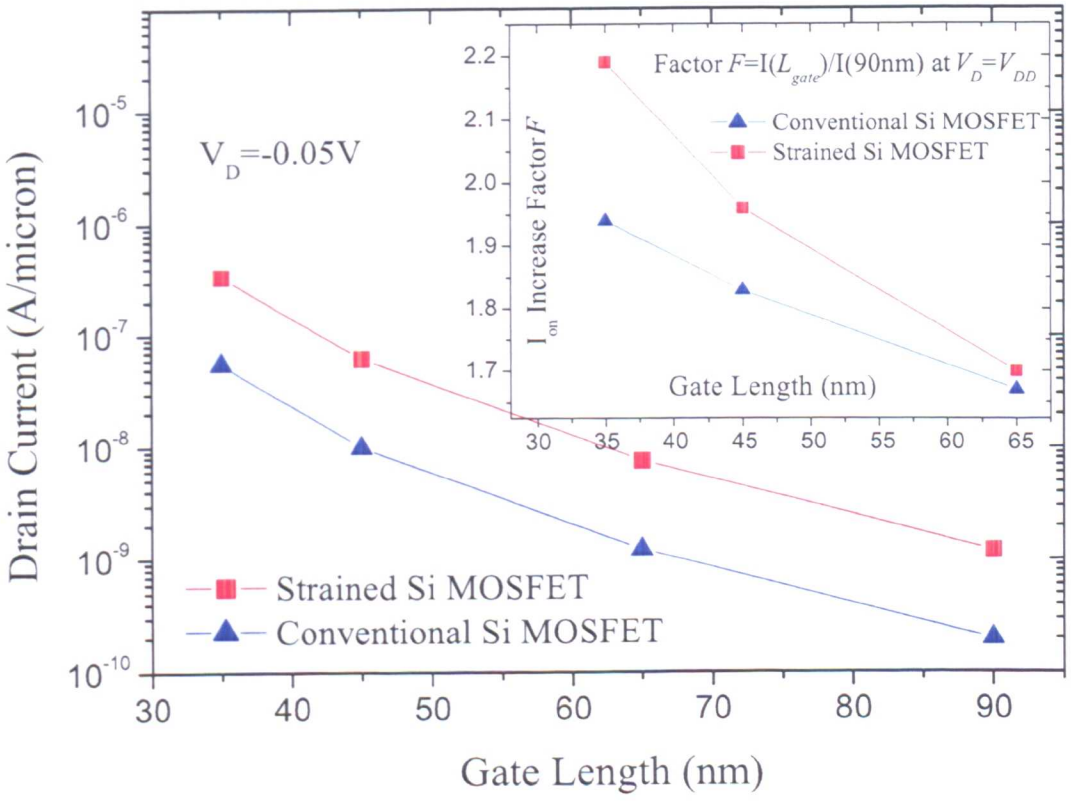


Figure 5.11: The low drain currents of Si control and strained Si MOSFETs as a function of gate length (scaling from 90 nm); the inset shows the high drain on current increase of scaled Si and strained Si MOSFETs as a function of gate length compared to the 90 nm devices

Fig. 5.11 plots the simulated device characteristics at low drain voltage and the scaling induced current increase at high drain bias as a function of gate length. It is seen from the figure that the high drain current increases during the scaling for both sets of devices and the current enhancement factor of strained Si MOSFETs

(over conventional Si MOSFETs) increases by 10% when the gate length is scaled from 90 nm to 35nm. This is due to appreciable velocity overshoot in the strained Si devices as a result of the increased hole relaxation time in strained Si, which is not observed in the results predicted in Table 5.4, where the DDM model is applied. Note that the strained Si layer thickness is kept unchanged since simulations show that variation of the strained Si layer thickness from 5 nm to 15 nm has negligible effects on the I_D -V_G characteristics.

In such device structures with a single retrograde channel doping, following the scaling rule, the channel doping and the source/drain doping concentration increase very rapidly. The scaling up of doping concentrations, which is limited by the solid solubility of corresponding dopants [165], increases the electric field and reduces the carrier mobility. As an alternative, the well-tempered MOSFET is a promising structure that enables scaling into decananometer regime with appropriate suppression of short channel effects [168]. Instead of using retrograde channel doping, as in the conventional MOSFET structure investigated so far, the well-tempered MOSFET adopts halo implants around the source/drain regions providing a raised source-to-drain barrier to suppress DIBL. The basic Si device structures simulated here are taken from the n-type 50 nm and 25 nm effective gate length MOSFETs by MIT [168], with appropriate modifications for the p-type devices. The halo dopings have been modified to obtain reasonable threshold voltages and a 10 nm strained Si layer on a relaxed $Si_{0.72}Ge_{0.28}$ substrate is used to realize the strained Si p-type well-tempered MOSFET, as illustrated in Fig. 5.12.

Table 5.5 presents the proposed device parameters and the simulated results of the well-tempered conventional Si and strained Si MOSFETs. As can be seen from the table, the proposed well-tempered devices offer comparable device performance improvement to the calibrated MOSFET structures studied beforehand. The subthreshold slope varies from 89 to 110 mV/decade and the threshold voltage roll off degrades to 200 mV in the 25 nm device. However, the off current of the 25 nm strained Si MOSFET (0.09 μA/μm) is slightly higher than the value (0.07μA/μm) set out in the ITRS roadmap [1]. When considering further device scaling, ad-

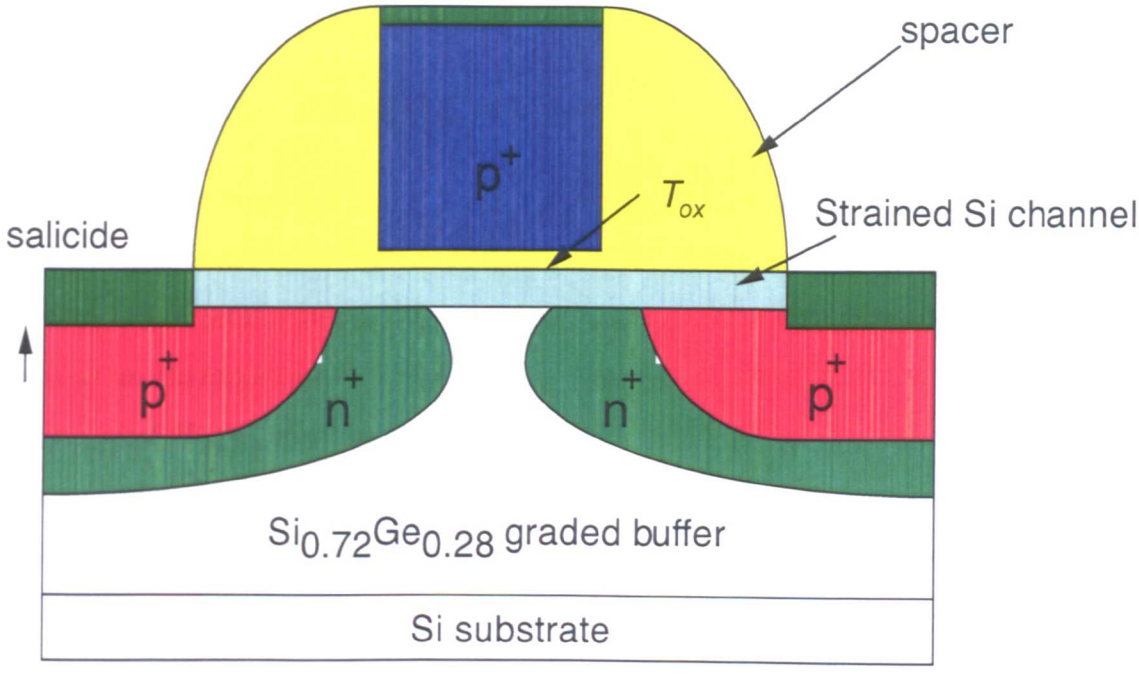


Figure 5.12: Schematic of the simulated p-type well-tempered strained Si MOSFET

ditional well and pocket engineering may be required in order to minimize short channel effects.

Table 5.5: The device parameters and the simulated results of well-tempered Si and strained Si MOSFETs

Effective gate length L_{eff} (nm)		50	25
Power supply voltage V_{DD} (V)		-1.2	-1.0
S/D peak doping N_A (cm^{-3})		2×10^{20}	2×10^{20}
Halo peak doping N_D (cm^{-3})		0.8×10^{19}	1.3×10^{19}
Gate oxide thickness t_{ox} (nm)		2.0	1.5
Junction depth x_j(nm)		37	32
Threshold voltage roll off ΔV_T (mV)	Si	90	205
	Strained Si	80	200
Subthreshold slope S (mV/decade) at V_D=V_{DD}	Si	91	109
	Strained Si	93	110
Off Current I_{off} (pA/μm) at V_D=V_{DD}	Si	10	1.3×10^4
	Strained Si	90	9×10^4
On Current I_{on} (A/m) at V_D=V_G=V_{DD}	Si	200	350
	Strained Si	300	570

5.4 Circuit Behaviour of Strained Si MOSFETs - 3-Stage Ring Oscillator

Based on the device scaling summarized in Tables 5.3 and 5.4, strained Si based CMOS circuits are expected to deliver enhanced performance as compared to conventional Si CMOS circuits. Here, a 3-stage ring oscillator (shown in Fig. 5.13) is used as an example to assess the strained Si CMOS circuit behaviour. In Fig. 5.13, W_n/L_n and W_p/L_p are the gate width/length ratios for the n- and p-type devices respectively; C_L is the load capacitance. For all circuits simulated here, W_n/W_p is set to be $1\mu m/2\mu m$ in order to balance the performance of the n-type and p-type MOSFETs.

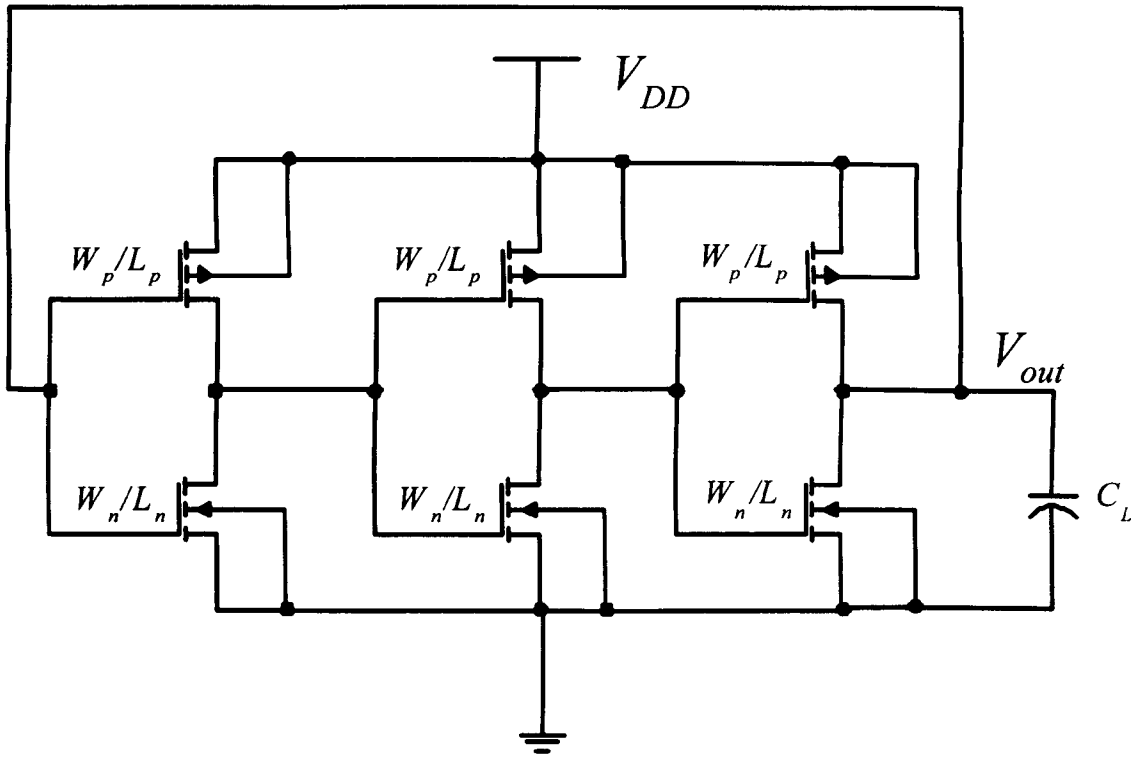


Figure 5.13: Schematic of CMOS 3-stage ring oscillator (W_n/W_p–$1\mu m/2\mu m$)

Compared to similar work on the performance assessment of strained Si CMOS using compact modelling [169], the drift-diffusion model based numerical modelling is used here to simulate circuits in order to directly obtain reliable information from discrete devices. Although this approach is more physical than the compact

modelling, it can only cope with small scale circuits and is not appropriate for practical circuit simulations. The circuit simulations carried out here are based on the device scaling study summarized in Tables 5.3 and 5.4. A MEDICI input file for the circuit shown in Fig. 5.13 may be created by describing the netlist file of the circuit and including all models directly from the device modelling for those discrete devices. By performing transient simulations, the MEDICI device simulator acts as a circuit simulator and outputs circuit characteristics. Fig. 5.14 shows the circuit characteristics of conventional Si and strained Si CMOS unloaded 3-stage ring oscillators. The n- and p-type devices are taken from the 67 nm effective gate length MOSFETs listed in Tables 5.3 and 5.4. The delay per stage in the strained Si circuit is about 6.5 ps less than that in the conventional Si circuit, showing a 19% circuit performance enhancement of the strained Si CMOS in such circuit structure without load capacitance.

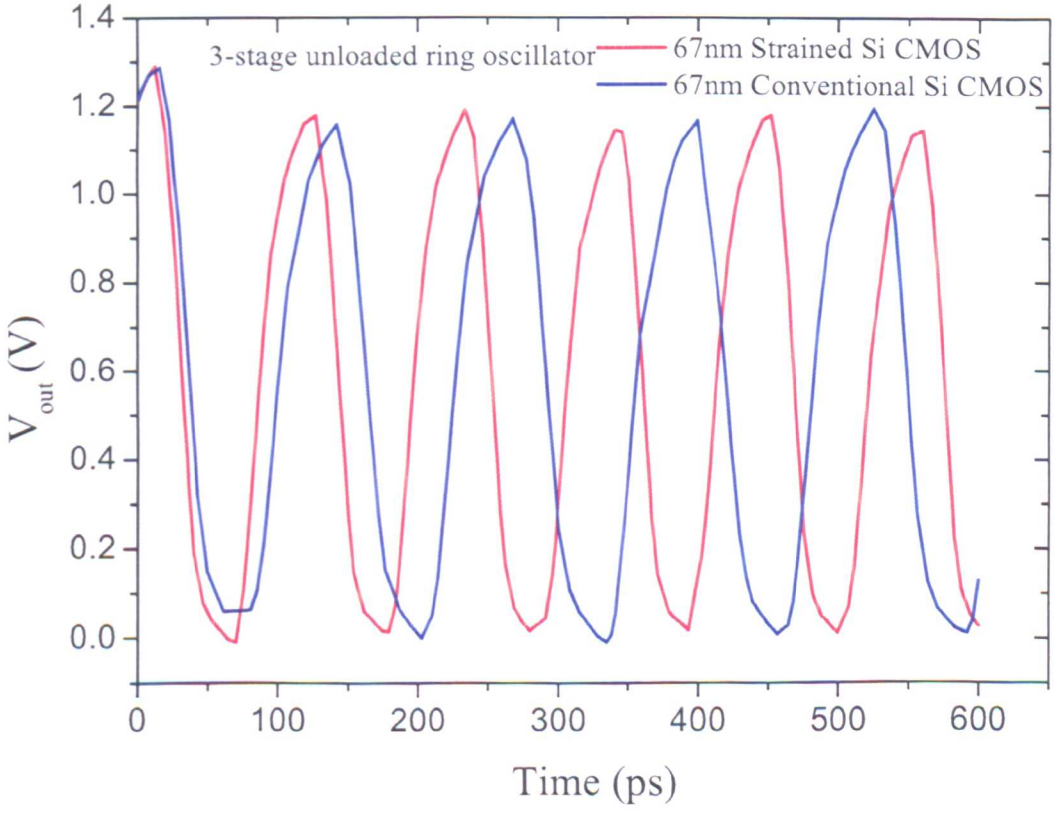


Figure 5.14: Output circuit characteristics for the 67 nm effective gate length Si and strained Si CMOS unloaded 3-stage ring oscillators

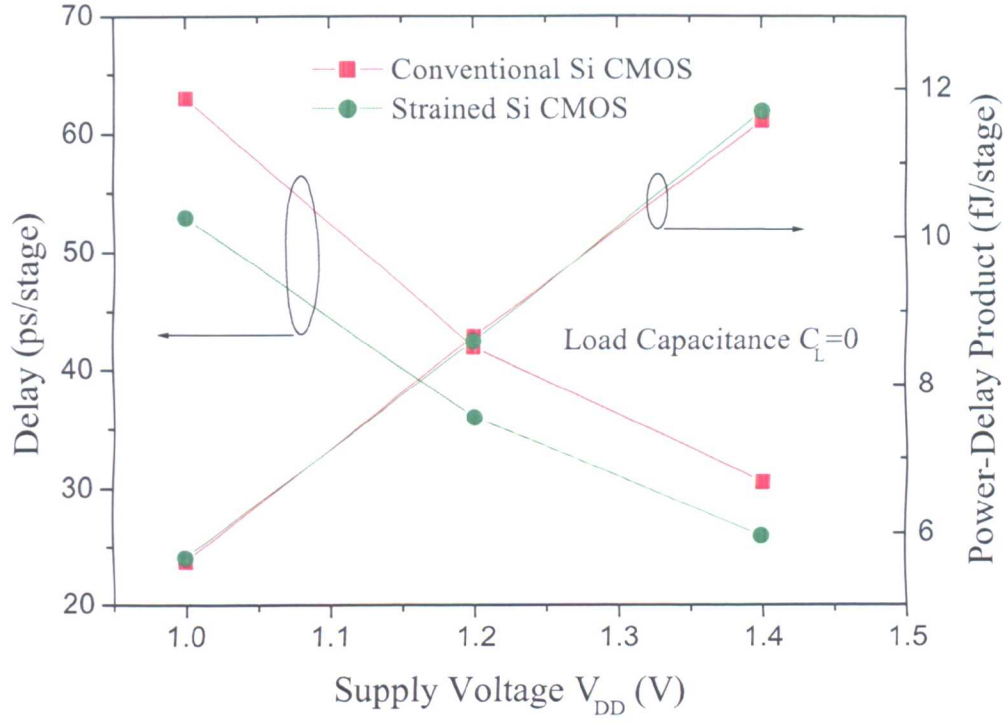


Figure 5.15: Circuit delays and power-delay products versus the supply voltage for the 67 nm Si and strained Si CMOS unloaded 3-stage ring oscillators

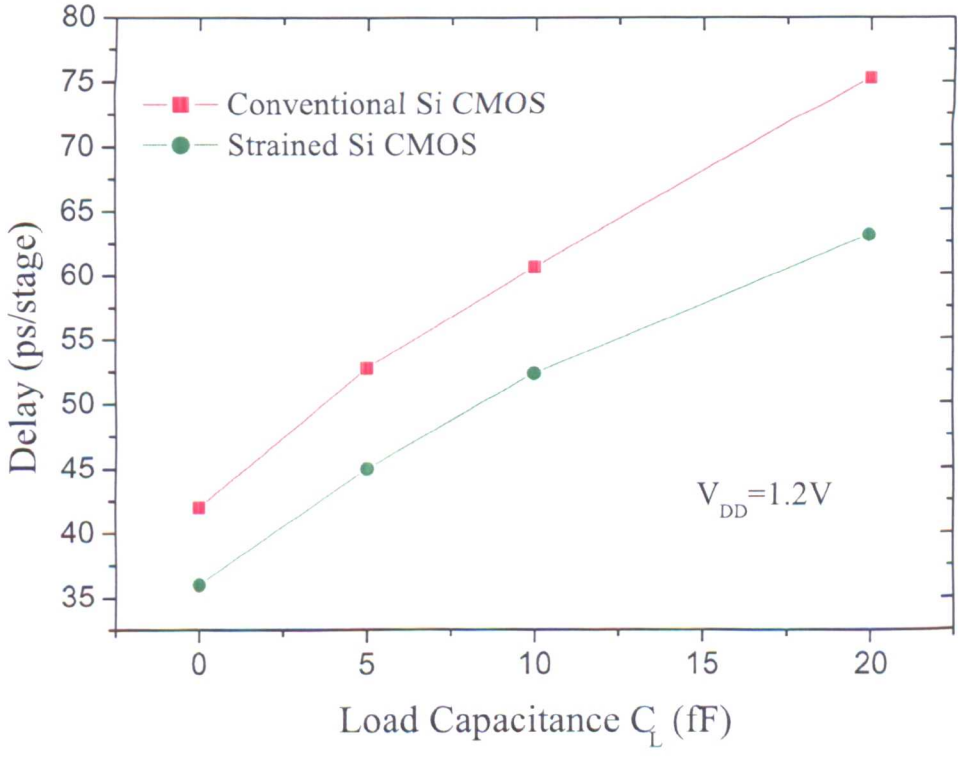


Figure 5.16: Circuit delays versus the load capacitance for the 67 nm Si and strained Si CMOS 3-stage ring oscillators

Fig. 5.15 compares the simulated circuit delays and power-delay products of the 67 nm conventional Si and strained Si ring oscillators when changing the power supply voltage V_{DD}. As V_{DD} decreases, the circuit delay of both circuits increase and the performance of the strained Si CMOS enhances slightly more. However, the power-delay products of both circuits remain unchanged when varying the power supply voltages. Fig. 5.16 plots the circuit delays for the 67 nm Si and strained Si ring oscillators as a function of load capacitance, C_L. The increase of the strained Si circuit delay is slower than that of the conventional Si circuit delay as the load capacitance increases, indicating that a performance enhancement of more than 19% for the strained Si circuit over the conventional Si circuit is expected in the real circuits with a realistic load capacitance.

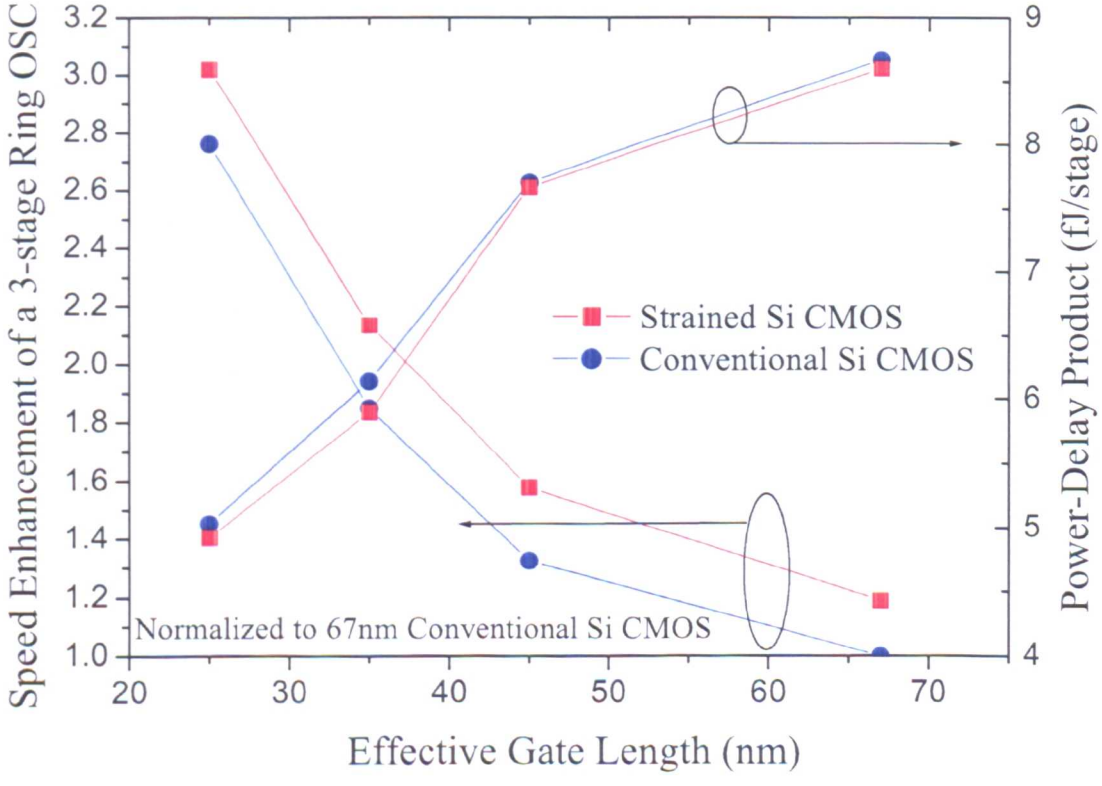


Figure 5.17: Speed enhancement and power-delay products against the effective gate length for Si and strained Si CMOS unloaded 3-stage ring oscillators

Using the scaled devices with 45 nm, 35 nm and 25 nm effective gate lengths listed in Tables 5.3 and 5.4, conventional Si and strained Si ring oscillators corresponding to

these different gate lengths are simulated. By normalizing the circuit delays of these ring oscillators to that of the 67 nm conventional Si circuit, the speed enhancement factors of these 3-stage ring oscillators are plotted in Fig. 5.17. It is evident that the circuit performance improves dramatically when the gate length is scaled down. The power-delay products of those circuits reduce as scaling down the gate length. Since the same scaling factors are used for both conventional Si and strained Si MOSFETs (see Tables 5.3 and 5.4), the circuit delays of both circuits scale by a factor of $1/\alpha\kappa$ to $1/\kappa$ and the power-delay products of both circuits change by a factor of α^2/κ^3, depending on the degree of velocity saturation [165], in general agreement with the trends shown in Fig. 5.17.

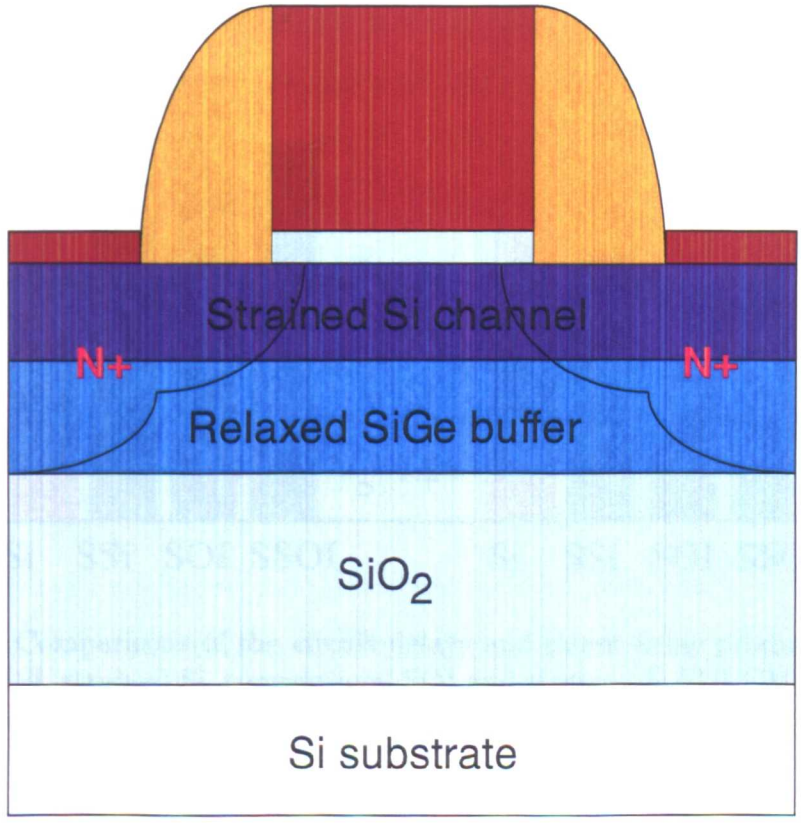


Figure 5.18: Illustration of simulated strained Si on SiGe on insulator device structure

To further improve the circuit performance, SOI device structures may be used in addition to the strained Si technology. Based on the 67 nm effective gate length devices listed in Tables 5.3 and 5.4, SiO_2 buried layers are used in the devices to

form an SOI structure. To make a simple test case, the body thicknesses (the distance from the Si/SiO_2 interface to the surface of the buried SiO_2 layer) of these conventional Si and strained Si SOI (SSOI) devices are assumed to be the same as the S/D contact junction depth (x_{j1} in Fig. 5.3). Fig. 5.18 illustrates the formed 67 nm strained Si on SiGe on insulator MOSFET device structure. The simulated 67 nm conventional Si SOI device structure also has buried SiO_2 layer in the same position as that in the SSOI device.

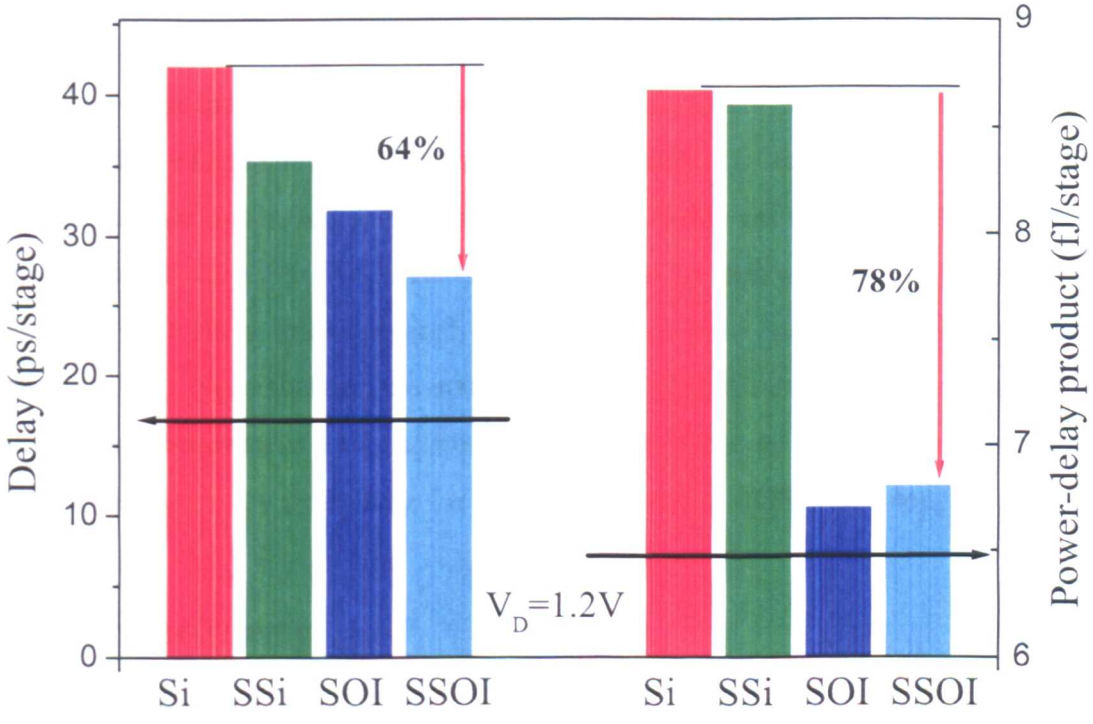


Figure 5.19: Comparisons of the circuit delays and power-delay products of 67 nm conventional Si, strained Si, conventional SOI and strained Si SOI CMOS unloaded 3-stage ring oscillators (W_n/W_p=1μm/2μm)

Fig. 5.19 compares the circuit delays and the power-delay products of the conventional bulk Si (Si), strained bulk Si (SSi), conventional SOI (SOI) and strained Si SOI (SSOI) unloaded 3-stage ring-oscillators. It is evident that the SOI circuits exhibit significant performance enhancements and much lower energy dissipations over their bulk counterparts. This is due to the reduced junction capacitances and better controlled short channel effects in the SOI structures [170]. The simulated

SOI devices are considered to be operating in a partially depleted (PD) model. In reality, optimizations are necessary in the SOI device and circuit designs in order to consider special properties of SOI devices [170], such as the floating body effect.

According to the 2003 edition of ITRS [1], several performance boosters, including transport-enhanced FETs (such as strained Si MOSFETs studied in this work), ultra-thin body SOI FETs, source/drain engineered FETs and multi-gate FETs, have been proposed and one or more than one of them may be required for devices beyond the 90 nm technology node in order to sustain the historic annual increase of intrinsic speed of high-performance MPUs at 17% [1]. Source/drain engineering, such as using metallic source and drain electrodes, is required for ultra scaled devices in order to maintain the source and drain resistance to be a reasonable fraction of the channel resistance [1]. Therefore, the scaled devices in this work may be further optimized to have smaller source/drain resistances and deliver faster circuit speed. The PD SOI structures simulated in this work may also be further modified to form a fully depleted (FD) SOI structure and enable extreme device scaling [170]. Nevertheless, the strained Si SOI CMOS, delivering remarkable performance enhancement and consuming low rate of power as compared to its bulk counterparts, is very promising for future advanced CMOS technology.

5.5 Summary

Comprehensive drift-diffusion and hydrodynamic simulations have been used to assess the device and circuit behaviours of sub-100 nm strained Si CMOS. The parasitic channel in the SiGe buffer is found to have negligible effects on the performance and leakage of sub-100 nm devices with high channel doping. Ge concentrations in the SiGe buffer in the range of 30%-40% could provide optimum device performance. The single retrograde channel doped MOSFET structure obtained by the scaling of published IBM devices shows a healthy performance improvement down to 35 nm physical gate length, keeping control of the short channel effects in the Si and the strained Si transistors. The well-tempered MOSFET structure is also appropriate for p-channel strained Si devices for gate lengths down to 25 nm. Increasingly com-

plicated well and pocket profile designs however are required for scaling the strained Si devices below the 20 nm barrier. 3-stage ring oscillators constructed by the scaled devices are simulated using the MEDICI device simulator. Strained Si circuits exhibit enhanced performance compared to their conventional Si counterparts. By incorporating the SOI device structure, the simulations show that strained Si SOI circuits deliver significant speed enhancement over all their competitors and exhibit reduced energy dissipation compared to their bulk counterparts.

This chapter predicts the device and circuit behaviours of the strained Si CMOS technology by the MEDICI simulator. However, as stated beforehand, the models used in this chapter cannot fully account for the real transport within short channel devices. Therefore, when considering non-equilibrium transport within the scaled devices, the performance enhancement of the strained Si CMOS are expected to be larger than the predictions made in this chapter. Nevertheless, the MEDICI simulator, which can simulate device behaviour under equilibrium conditions, still provides useful information for the simulated devices and circuits. To properly estimate the device performance of sub-100 nm strained Si MOSFETs, more advanced simulations are required, such as Monte Carlo. These are considered in the next chapter.

Chapter 6

Performance Predictions of sub-100 nm n-type Strained Si MOSFETs

As discussed in Chapter 3, the DDM and HDM simulation models adopted in this study are based on different degrees of approximation of the moments of the Boltzmann transport equation. Such treatments enable efficient numerical device simulations and are widely used to assess device behaviour during the device design phase. However, as devices are scaled into the sub-100 nm regime, the appearance of non-equilibrium carrier transport requires more precise solutions of the Boltzmann transport equation in order to capture the physics of carrier transport when subjected to very high electric fields. The Monte Carlo (MC) method is the most widely used approach to solve the Boltzmann transport equation. An ensemble Monte Carlo simulator is used in this chapter to simulate sub-100 nm n-type strained Si MOSFETs and assess device performance. The subthreshold behaviour of these devices, studied by the DDM and HDM simulations in Chapter 5, are not the interest of this chapter since the Monte Carlo method is a statistical approach and may induce large numerical noise at low electric field (subthreshold region). The device structures simulated in this chapter are based on the 80 nm gate length conventional Si and strained Si MOSFETs which were calibrated in Chapter 5. The device structures and doping profiles used in the Monte Carlo simulations are directly taken from MEDICI DDM simulations. Successful calibration of devices from Drift-Diffusion simulations ensure reliable device information for the Monte Carlo simulation.

The transport enhancement of electrons and holes in strained Si derives from the lower effective masses and from the reduction in the inter-valley phonon scattering due to the strain induced X-valley splitting. Such mobility enhancements in strained Si have been simulated by Monte Carlo and are plotted in Figs. 2.3 and 2.4 (see Chapter 2) for electrons and holes respectively. However, to fully evaluate MOSFET device performance, it is also necessary to incorporate the effect of interface roughness (IR) scattering. Interface roughness has long been an important scattering mechanism limiting the device performance of surface channel MOSFETs [171]. In modern sub-100 nm devices, the high channel doping induces a high perpendicular electric field, reinforcing the degree of interface scattering and resulting in an undesired degradation in channel mobility and a corresponding reduction in device current. It is therefore necessary to study interface roughness in the strained Si surface channel MOSFET and compare it with the conventional Si MOSFET in order to understand the device performance enhancement mechanisms in the strained Si MOSFETs. This chapter uses the interface roughness model [36] developed within the Device Modelling Group at the University of Glasgow and validates the model by calibrating it with respect to the universal mobility behaviour [171] and device characteristics [31] of conventional Si MOSFETs. Based on this understanding, Monte Carlo simulations are then carried out to study the impact of interface roughness on the performance enhancement of strained Si MOSFETs and predict the performance of scaled strained Si MOSFETs down to a 35 nm gate length.

When the gate length is down to 35 nm, gate oxides as thin as 1.1 nm may be required, as indicated in Chapter 5 (see Table 5.3). Such thin gate oxides results in an intolerably high gate leakage. A possible solution to this problem is the incorporation of high-κ dielectrics in the gate stack [1]. However, while the integration of high-κ dielectrics may solve the gate leakage problem, it introduces a number of technological problems: inducing low thermal stability; a high density of interface states and fixed charges; and a poor quality of the top and bottom interfaces [5]. Apart from these technological issues, a fundamental drawback associated with the introduction of high-κ dielectrics in the gate stack is the mobility degradation due to strong soft-optical (SO) phonon scattering. This scattering arises from the cou-

pling between carriers in the inversion layer and surface longitudinal optical (LO) phonons at the silicon/high-κ interface due to the increased polarizability of high-κ dielectrics [6]. The effect of such scattering has been demonstrated experimentally in the case of HfO_2 [7, 96] and supported by a temperature dependence study of mobility [98]. However, there is very little quantitative analysis of the impact of this scattering mechanism on device performance. This chapter studies the impact of high-κ dielectrics on sub-100 nm conventional and strained Si n-type MOSFETs by including SO phonon scattering into Monte Carlo device simulations. The effect of the two leading high-κ gate dielectrics, HfO_2 and Al_2O_3, has been examined.

6.1 Simulation Methodology

6.1.1 Monte Carlo simulations

Monte Carlo [172] is a statistical numerical method used for solving mathematical problems. When applied to carrier transport in solids, Monte Carlo can, in theory, provide an precise numerical solution of the BTE without the necessity of solving it directly. The essence of Monte Carlo is to simulate the motion of the ensemble of carriers in **k** space as well as in **r** (real) space. The motion of each carrier is governed by semiclassical equations of motion and by stochastic collisions with various perturbations (phonons, ions). A typical Monte Carlo process flow for the simulations of a stationary and homogeneous transport process is illustrated in Fig. 6.1 [172]. The starting point of the program is the definition of the physical system of interest, including the parameters of the material and the values of physical quantities. The parameters controlling the simulation, such as the duration of each sub-history and the desired precision of the results, are also defined at this level. The next step in the program is a preliminary calculation of each scattering rate as a function of electron energy, providing information on the maximum value of these functions for optimizing the efficiency of the simulations.

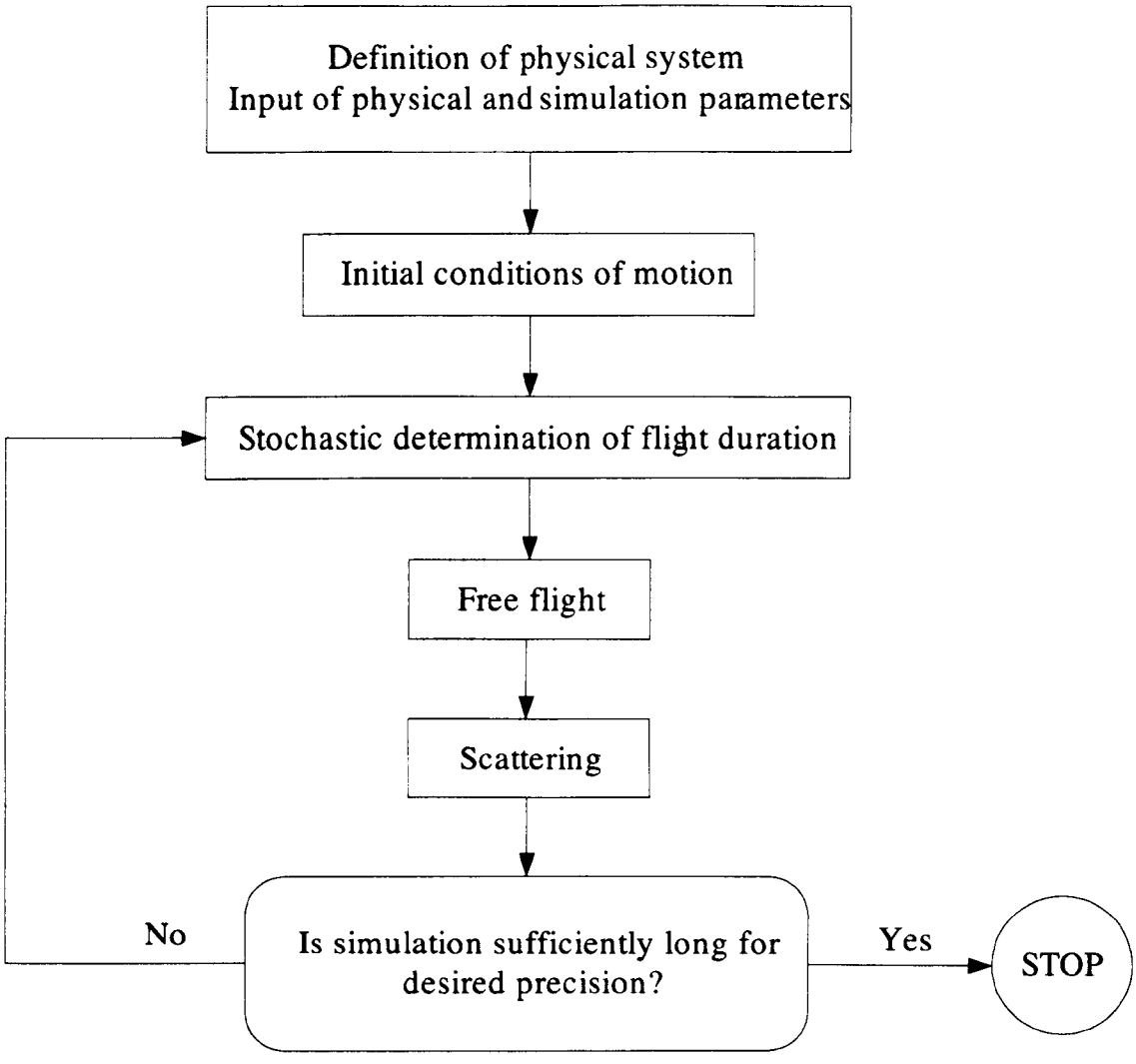


Figure 6.1: Flowchart of a typical Monte Carlo program

In the case under consideration, in which a steady-state situation is simulated, the time of simulation must be long enough so that the initial conditions of the carrier motion do not influence the final results. Generally, the longer the simulation time, the less influence the initial conditions will have on the average results. The choice of a "good" time of simulation is a compromise between the need for ergodicity (requiring an infinite simulation time) and need to be computationally efficient.

The subsequent step in the program is the generation of the flight duration. The electron wave vector $\mathbf{k}$ changes continuously during free flight because of the applied electric field. Thus if $\lambda[\mathbf{k}(t)]dt$ is the probability that an electron in the state $\mathbf{k}$ suffers a collision during the time dt, the probability $P(t)$ that the electron will suffer its next collision during dt around t is given by [172]

$$P(t)dt = \lambda[\mathbf{k}(t)] \exp\left[-\int_0^t \lambda[\mathbf{k}(t')]dt'\right] dt \tag{6.1}$$

Because of the complexity of the integral in the exponent, the concept of "self-scattering" [172] is introduced in order to overcome the difficulty of using evenly distributed random numbers to generate stochastic free flights with the distributions of equation (6.1). Once the electron free flight is terminated, the scattering mechanism has to be selected. The weight of the i-th scattering mechanism is given by [172, 173]

$$P_i(\mathbf{k}) = \frac{\lambda_i(\mathbf{k})}{\Gamma}, \qquad \Gamma = \sum_{i=1}^{n} \lambda_i(\mathbf{k}) \tag{6.2}$$

A scattering mechanism is selected by generating random number r between 0 and 1 and by testing the inequalities [172, 173]

$$\sum_{i=1}^{j-1} \frac{\lambda_i(\mathbf{k})}{\Gamma} < r < \sum_{i=1}^{j} \frac{\lambda_i(\mathbf{k})}{\Gamma}, \qquad j = 1, \cdots, n \tag{6.3}$$

The j-th scattering mechanism is selected if the j-th inequality is fulfilled.

The next step in the program is the choice of state after scattering. If the free flight ended with a self-scattering, the new state after scattering event, $\mathbf{k}_f$, is equal to the state before scattering, $\mathbf{k}_i$. When a true scattering event occurs, $\mathbf{k}_f$ must be generated stochastically according to the differential cross section of that particular scattering mechanism and conservation of momentum and energy. The last step of the simulation is the collection of statistical averages.

This chapter uses an ensemble Monte Carlo simulator [29, 43] developed within the Device Modelling Group. The Si band-structure for electrons employed in the Monte Carlo simulator consists of a set of 6 non-parabolic ellipsoidal Δ-valleys. The six conduction band valleys are included through three pairs: valley pair 1 pointing in the <100> direction; valley pair 2 in the <010> direction and valley pair 3 in the <001> direction. The Monte Carlo simulator includes all the relevant scat-

tering mechanisms [29, 43]: inelastic acoustic phonon scattering, ionized impurity scattering, f- and g- optical intervalley phonon scattering, interface roughness scattering and soft-optical phonon scattering. All of these scattering mechanisms are important for being able to correctly model electronic transport in strained Si and SiGe layers. Interface roughness scattering and soft optical phonon scattering are discussed in the following sections. Using this simulator, the electron and hole mobilities in strained Si as a function of Ge concentration in SiGe substrate have been calculated as plotted in Figs. 2.3 and 2.4.

6.1.2 Interface Roughness Scattering

To date, two main approaches which have been adopted to account for interface roughness scattering within the ensemble Monte Carlo framework. The first uses a phenomenological parameter corresponding to the fraction of diffuse and/or specular reflections, which are independent of both the energy and the incident angle of carriers [174, 175]. However, this approach lacks physical insight and the results are sensitive to the arbitrary factors for specular and diffuse scattering. The second is based on the fluctuations of the eigen-states of the carriers within the inversion layer [176] but is however expensive at present for efficient device simulation. This approach is also unsuitable considering the semi-classical nature of the Monte Carlo formalism.

This work uses an interface scattering model developed in the Device Modelling Group [36] which calculates the probabilities of specular and diffuse scattering as a function of the physical parameters associated with both the carriers and the interface. The model treats the interface roughness scattering non-perturbatively by incorporating the effects of scattering as a boundary condition for the Boltzmann Transport Equation. The model [36] is valid at large distances from the object plane, r. In the case of interface roughness scattering the assumptions are fulfilled when $r > \Delta^2/\Lambda_{dB}$, where Λ_{dB} is the electron de-Broglie wavelength and Δ is the RMS height of the interface above the plane. In Si MOSFETs with SiO_2 gate dielectrics, the thermal DeBroglie wavelength Λ_{dB} at room temperature is about 4.3 nm [177] and Δ is typically less than 1 nm. Therefore the model is applicable for

surface channel MOSFETs since $r > \Delta^2/\Lambda_{dB}$ is satisfied for most carriers in the inversion layer. The model treats the incident carrier as a plane wave from which the emergent flux (which is interpreted as probability using quantum mechanics) is estimated based on the principles of the geometrical scattering from a perfectly reflecting surface. This method is suitable for those surfaces where $\lambda \gg \Delta$, where λ is the correlation length of the roughness distribution. The model neglects the potential variations within the semiconductor in the region of the interface; the assumption is that the carriers are unconfined. This model therefore holds for high-energy electrons, such as those that occur in modern nano-scale devices at room temperature, especially in the on-current regime.

The model assumes an exponential autocorrelation function [178] to define the Si/SiO_2 interface and has been used to calculate the probability of specular scattering, P_F, and the angular scattering probability for diffuse scattering $P_{Diffuse}$ [36]. The probability for specular scattering is given by:

$$P_F = \exp(-4\Delta^2 k^2 \cos\theta_{IN}), \tag{6.4}$$

where k represents the k-vector of the scattering particle and θ_{IN} represents the incident angle of the scattering particle as measured from the normal. In the case of a diffuse scattering event, the angular probability distribution, $P_{Diffuse}$, is given by:

$$P_{Diffuse} \propto \frac{(\cos\theta_{IN} + \cos\theta_{OUT})^2}{1 + \frac{\lambda^2 k^2 (\sin\theta_{IN} + \sin\theta_{OUT})^2}{2}}, \tag{6.5}$$

where θ_{OUT} represents the emergent angle of the scattering particle. It is clear that the probabilities are now related to physical parameters of Δ and λ. The probability of specular reflection increases as the interface becomes smoother $\Delta \rightarrow 0$ and the angular probability for diffuse scattering forms a lobe around the specular scattering angle, becoming increasingly diffuse and spread out as $\lambda \rightarrow 0$.

6.1.3 Soft-optical Phonon Scattering

The introduction of high-κ gate dielectrics reduces the gate leakage current typically by orders of magnitude [1]. However, it also introduces strong soft optical phonon scattering. The scattering results from the strong ionic polarizability of the high-κ material, which also determines the large value of the dielectric constant. Electrons scatter from these phonons via a Fröhlich interaction, which has an unscreened scattering field amplitude given by [6]:

$$\phi = \left\{ \frac{\hbar\omega_{SO}}{2q^2} \left[\frac{1}{\epsilon_{Si}^{\infty} + \epsilon_{ox}^{\infty}} - \frac{1}{\epsilon_{Si}^{\infty} + \epsilon_{ox}^{0}} \right] \right\}^{1/2} \tag{6.6}$$

where ϵ_{ox}^{∞}, $\epsilon_{ox}^{0}$and ϵ_{Si}^{∞} are the optical and static permittivities for the oxide and Si respectively. ω_{SO} is the soft-optical phonon energy and is calculated from the two dominant transverse-optical (TO) phonon modes in the dielectric via the Lyddane-Sachs-Teller relationship [6]:

$$\omega_{SO} = \sqrt{\frac{1}{2a}\left[b \pm (b^2 - 4ac)^{1/2}\right]} \tag{6.7}$$

$$a = \epsilon_{ox}^{\infty} + \epsilon_{Si}^{\infty} \tag{6.8}$$

$$b = a(\omega_{TO1}^2 + \omega_{TO2}^2) + (\epsilon_{ox}^{i} - \epsilon_{ox}^{\infty})\omega_{TO2}^2 + (\epsilon_{ox}^{0} - \epsilon_{ox}^{i})\omega_{TO1}^2 \tag{6.9}$$

$$c = (a + \epsilon_{ox}^{0} - \epsilon_{ox}^{\infty})\omega_{TO1}^2\omega_{TO2}^2 \tag{6.10}$$

where ω_{TO1} and ω_{TO2} are the phonon energies for two TO modes (assuming $\omega_{TO1}<\omega_{TO2}$); ϵ_{ox}^{i} is the intermediate insulator permittivity describing the dielectric response of the insulator at some intermediate frequency between ω_{TO1} and ω_{TO2}. Equation (6.7) for calculating the SO phonon energy is based on the assumption of $Qt \rightarrow \infty$. t is the distance from the interface and Q is the in-plane momentum transfer. The scattering field ϕ whose amplitude is given by equation (6.6) falls off exponentially with distance t as $\exp(-Qt)$. Table 6.1 lists the static, intermediate and optical permittivities (relative to the vacuum permittivity ϵ_0), along with the phonon energies for two TO modes, and for the different dielectrics considered in this work.

Table 6.1: Parameters used to calculate the electron soft-optical phonon coupling in high-κ gate dielectrics, after Fischetti *et al.* [6]

Quantity/Dielectric	SiO_2	Al_2O_3	HfO_2
ϵ^0_{ox} (ϵ_0)	3.90	12.53	22.00
$\epsilon^i_{ox}(\epsilon_0)$	3.05	7.27	6.58
$\epsilon^\infty_{ox}(\epsilon_0)$	2.50	3.20	5.03
ω_{TO1} (meV)	55.60	48.18	12.40
ω_{TO2} (meV)	138.10	71.41	48.35
ω_{SO1} (meV)	57.14	53.19	16.79
ω_{SO2} (meV)	140.78	82.48	50.67

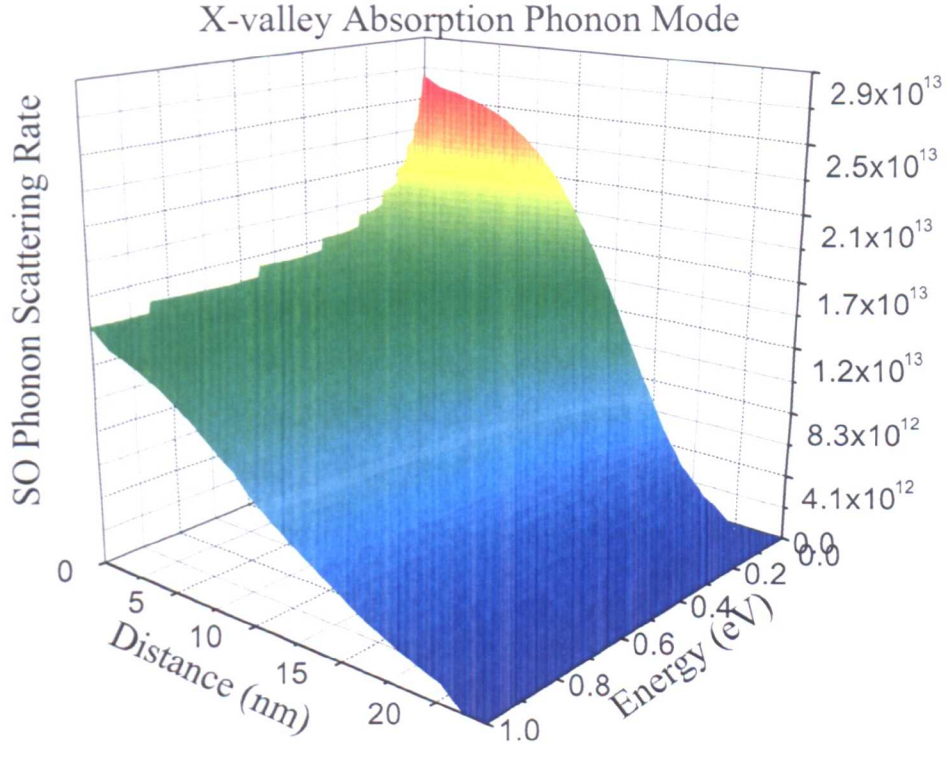


Figure 6.2: Scattering rate of soft optical phonon scattering in a Si/HfO_2 system (absorption mode of phonon mode 1 (ω_{SO1}) in X-valley)

It can be seen from Table 6.1 and equation (6.6) that for low phonon energies, the larger difference between the static and optical permittivities for the high-κ dielectrics, HfO_2 and Al_2O_3, causes strong scattering to carriers in the inversion layer. The high phonon energies (hard Si-O bond) and the small difference between the static and optical permittivities for SiO_2 make such scattering negligible in the

Si/SiO_2 based system.

The scattering rate for absorption of phonon mode 1 (ω_{SO1}) in the X-valley of Si is illustrated for Si/HfO_2 interface in Fig. 6.2. The scattering rate decreases with an increase in energy, which is characteristic of a Fröhlich type interaction and drops exponentially as a function of the distance from the Si/HfO_2 interface.

6.2 Interface Roughness Scattering in Strained Si MOSFETs

6.2.1 Model Validation

The interface roughness model has been implemented in the in-house ensemble Monte Carlo simulator and convincingly reproduces the universal mobility data for Si [171] (shown in Fig. 6.3) using Δ–0.5 nm and λ–1.8 nm, which is in good agreement with the experimental data (Δ–0.48 nm and λ–1.3 nm) of Goodnick *et al.* [178]. The measured mobility behaviour in the 67 nm effective gate length Si MOSFET calibrated in Chapter 5 [31] is also plotted for comparison. The effective mobility behaviour illustrated in Fig. 6.3 is extracted from a 1×1 μm^2 MOS device at low lateral field (0.1 $kV{\cdot}cm^{-1}$) with an uniform p-type substrate doping of 7.2×10^{16} cm^{-3}.

Using the calibrated roughness parameters, Δ–0.5 nm and λ–1.8 nm, Monte Carlo simulations are carried out to simulate the 67 nm effective gate length conventional Si MOSFET calibrated in Chapter 5. The device structure and the doping profiles are obtained after successful MEDICI calibrations and a summary of the calibrated device information is listed in Table 5.1. The device information from MEDICI together with the calibrated roughness parameters are then used in Monte Carlo simulations. In this way, we are able to use Monte Carlo simulation to reproduce the experimental I_D-V_G characteristics without further parameter calibration, as shown in Fig. 6.4. The drain current at low drain voltage without the IR scattering is also plotted as a reference. It is observed that interface roughness scattering degrades the drain current at low drain bias by ~35% at V_G-V_T–1.1 V.

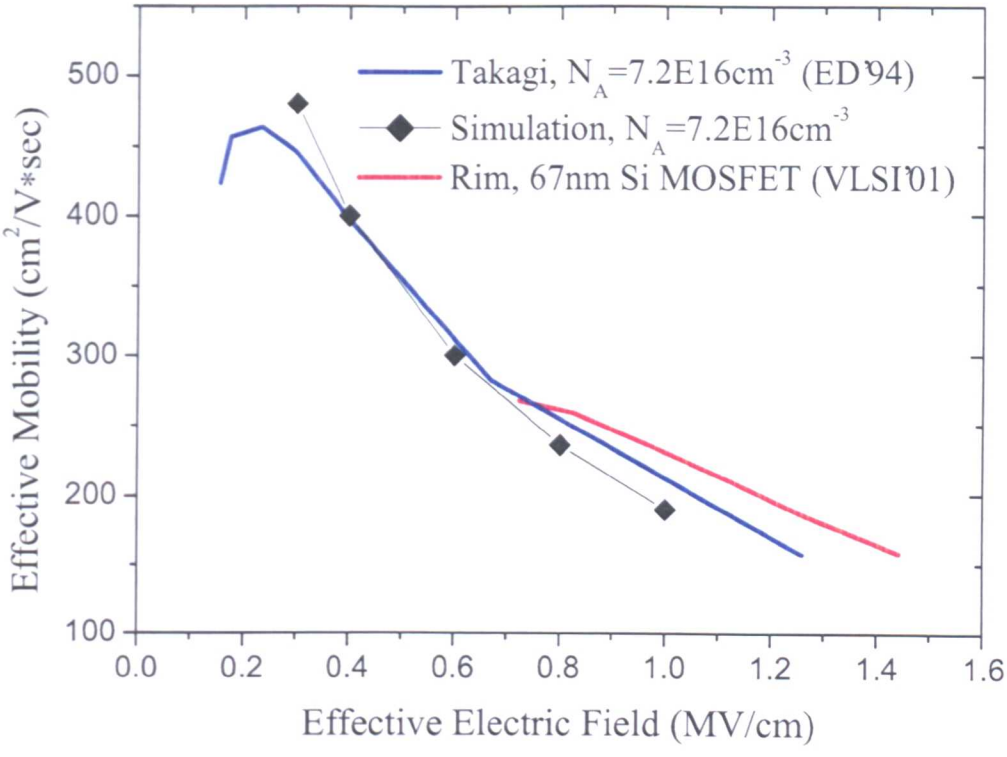


Figure 6.3: Field dependence of the effective electron mobility of bulk Si, compared with data from Takagi *et al.* [171] and Rim *et al.* [31]

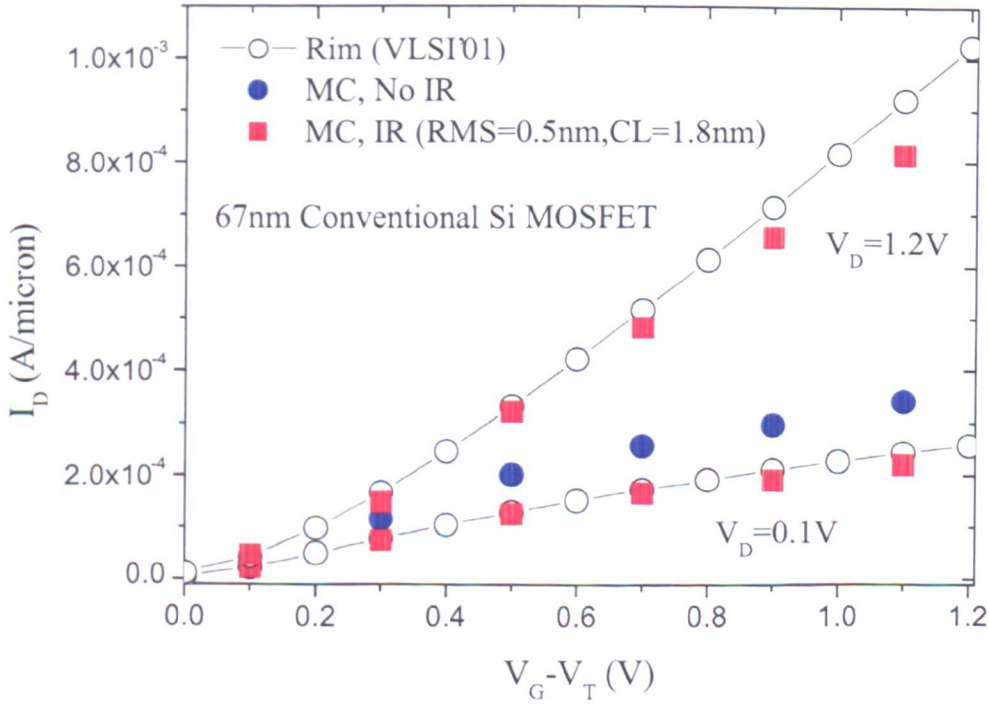


Figure 6.4: Monte Carlo simulated I_D-V_G characteristics of the 67 nm Si MOSFET with and without IR scattering, compared with data from Rim *et al.* [31]

6.2.2 Interface Roughness Scattering in Strained Si MOSFETs

The success of the interface roughness model in the case of relaxed Si suggests it may be suitable for the application to strained Si interfaces. To date, little work has been done on the interface roughness scattering strained Si MOSFETs, which is crucial in exploiting the performance enhancement of strained Si MOSFETs. A recent theoretical study by Fischetti *et al.* [179] suggested that the assumption of *an increasingly smoother interface with increasing (tensile) strain* may in part explain the observed performance in strained Si MOSFETs. This is in addition to the lower conductivity mass and the reduction of intervalley scattering within strained Si MOSFETs. This is supported by the fact that like the tensile strain in a strained Si MOSFET, *a sufficiently large confinement potential* within the inversion layer of a conventional Si MOSFET also generates a significant splitting between the doubly degenerate and fourfold degenerate subbands, and leads to significant reduction of the intervalley scattering and large electron population in the low conductivity mass subbands. In their work, the correlation length (CL) was chosen to be fixed and the RMS height of the surface to vary. In this way, they were able to reproduce the experimentally observed mobility behaviour of the strained Si MOSFET.

In this work, however, it is assumed that tensile strain changes only the correlation length of the roughness, the RMS height being unaffected [180]. Applying the interface roughness model to a strained Si layer grown on a relaxed $Si_{0.85}Ge_{0.15}$ buffer, the roughness parameters of unstrained Si (Δ=0.5 nm and λ=1.8 nm) lead to a lower effective mobility at high fields compared to the experimentally measured mobility data of the 67 nm effective gate length IBM strained Si MOSFET calibrated in Chapter 5, as shown in Fig. 6.5. This suggests an overestimation of the degree of interface roughness scattering in the strained Si layer. In order to reproduce the experimental data it is necessary to consider a *smoother* surface for the strained Si interface by increasing the correlation length, as shown in Fig. 6.5. It is determined that a correlation length, λ, of 3.0 nm gives the best agreement with experiment in the high-field regime, where interface roughness scattering dominates.

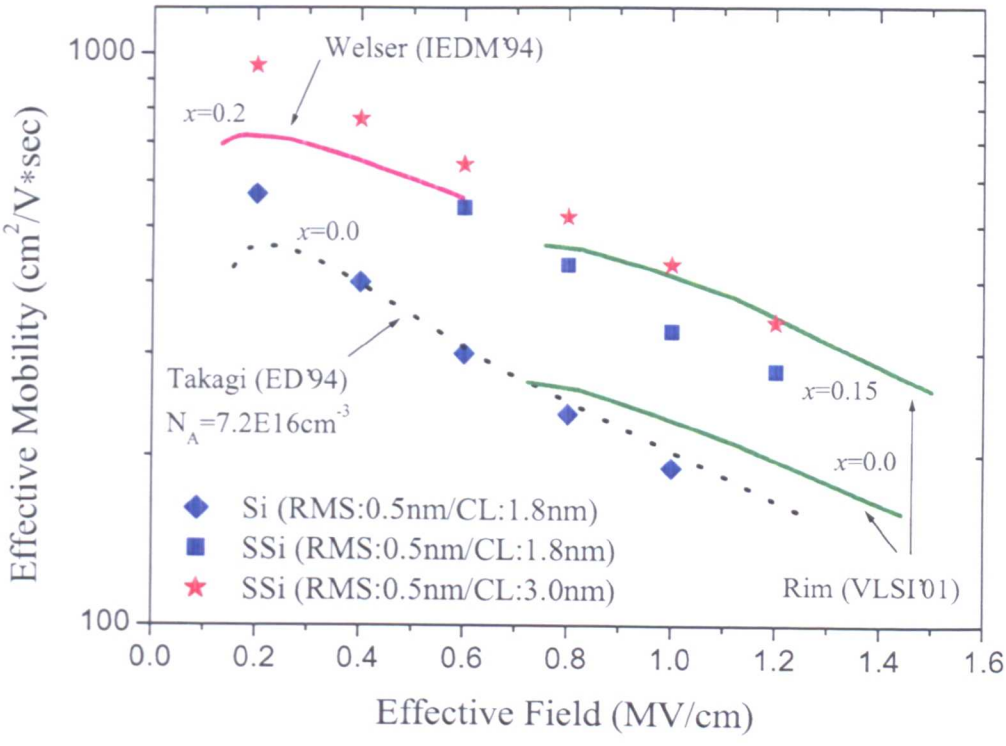


Figure 6.5: Field dependence of the effective electron mobility of strained Si, compared with data from Rim *et al.* [31] and Welser *et al.* [67]

The calibrated interface roughness parameters of Δ=0.5 nm and λ=3.0 nm are then used in Monte Carlo simulations of the 67 nm effective gate length IBM strained Si MOSFET. The device dimensions and doping profiles, summarized in Table 5.1, provide reliable information for the Monte Carlo simulations. Using the IR parameters of Δ=0.5 nm and λ=3.0 nm, the results of the Monte Carlo simulations are shown in Fig. 6.6. The simulations agree with the experimental data without any further adjustment of the roughness parameters. To assess the effect of the interface roughness scattering, drain currents at low drain voltage (V_D=0.1 V) without interface roughness scattering and with interface roughness scattering using the conventional Si (Δ=0.5 nm and λ=1.8 nm) parameters are also plotted for comparison. It is evident that using the conventional Si roughness parameters underestimates the drain current. From Fig. 6.6, the observed drain current degradation at V_G-V_T=1.1 V and V_D=0.1 V is about 25%, less than the 35% observed in the conventional Si

MOSFET as shown in Fig. 6.4, indicating that interface roughness scattering is playing a less important role within the strained Si MOSFET. The agreements between the simulations and the experimental mobility and device characteristics of the strained Si MOSFET adds further weight to the conjecture that a *smoother* interface in the case of strained Si (represented by a larger correlation length) plays a significant contribution in the performance improvement in strained devices.

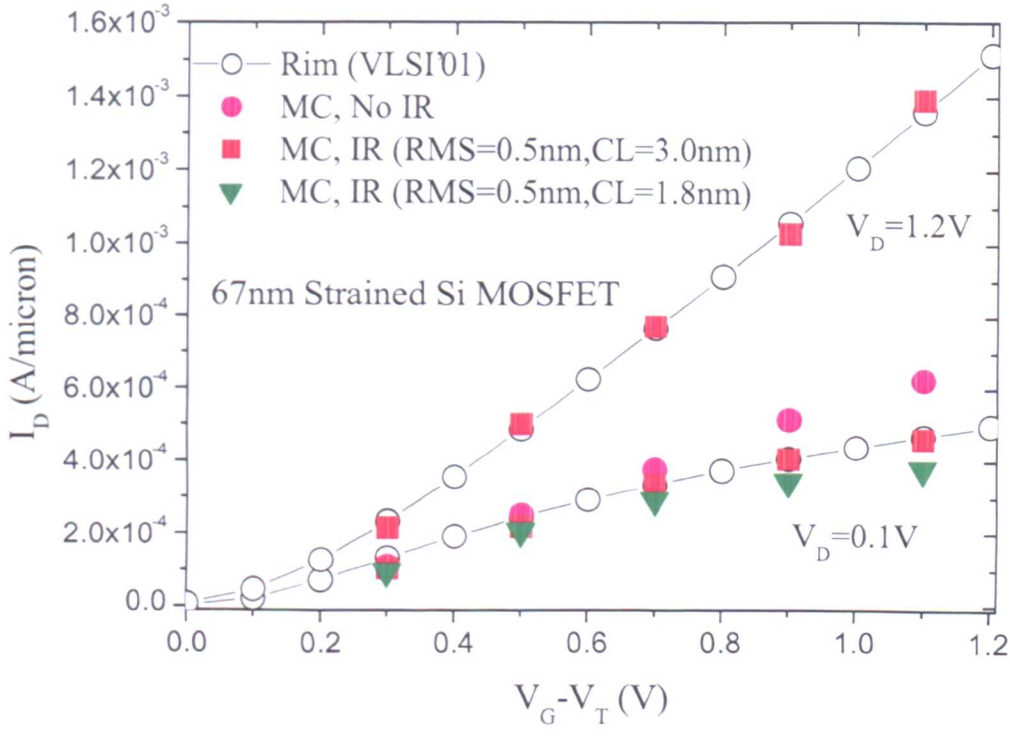


Figure 6.6: Monte Carlo simulator I_D-V_G characteristics of the 67 nm strained Si MOSFET with and without IR scattering, compared with data from Rim *et al.* [31]

Fig. 6.7 shows the average channel carrier velocities obtained from the Monte Carlo simulations in bulk Si and strained Si with and without interface roughness scattering. It is evident from the figure that the interface roughness scattering significantly affects carrier transport within the channel, limiting device performance at high fields. The average channel velocity in the strained Si device (using the calibrated roughness parameters Δ–0.5 nm and λ–3.0 nm) is higher than that

in the bulk Si device through whole channel region, contributing to the observed performance enhancement in the strained Si MOSFET.

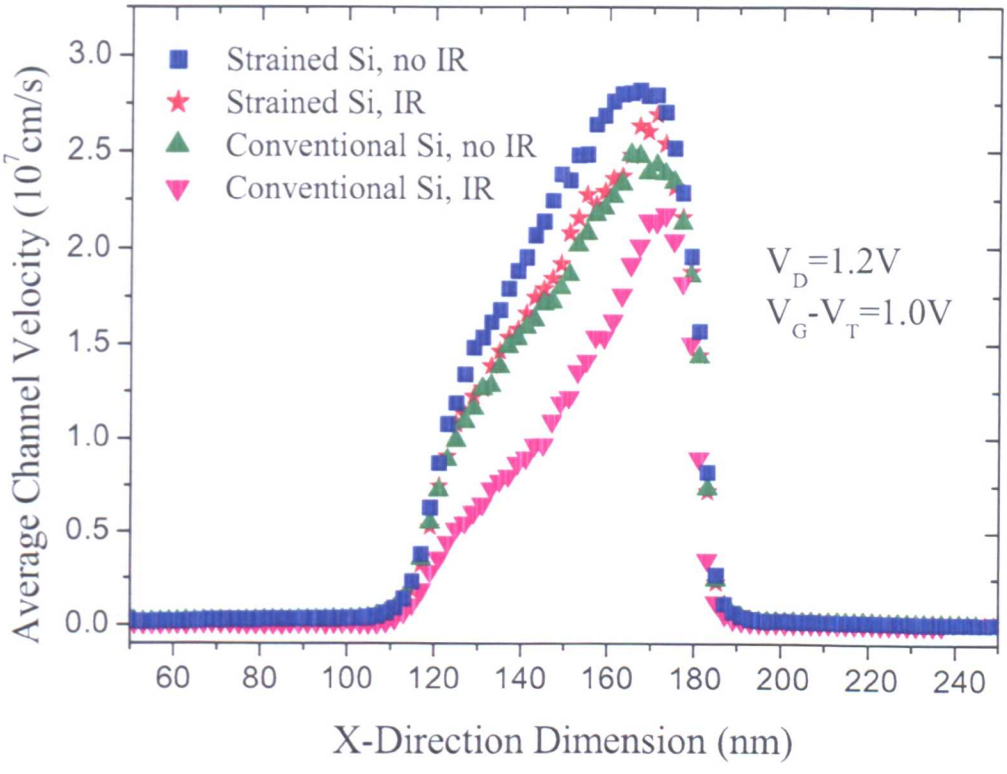


Figure 6.7: Monte Carlo simulated average channel velocities of the 67 nm n-type conventional Si and strained Si MOSFETs with and without interface roughness scattering

6.2.3 Performance Predictions of Scaled Strained Si MOSFETs

The validation and calibration of the interface roughness scattering model with the experimental data for both conventional Si and strained Si MOSFETs enable us to evaluate the performance of scaled strained Si MOSFETs. To ensure credibility, an n-type Si MOSFET fabricated by Toshiba [181] is used as the calibration standard: This device has a 35 nm physical gate length and 1.2 nm thick gate oxide (NO oxynitride) featuring relative permittivity of 4.75. The complete doping profiles are

obtained from calibrated Taurus [145] process and device simulations. The Taurus calibrations [1], which incorporate the same mobility models as that used in the MEDICI calibrations discussed earlier in Chapter 5, reproduce the experimental device characteristics, as shown in Fig. 6.8.

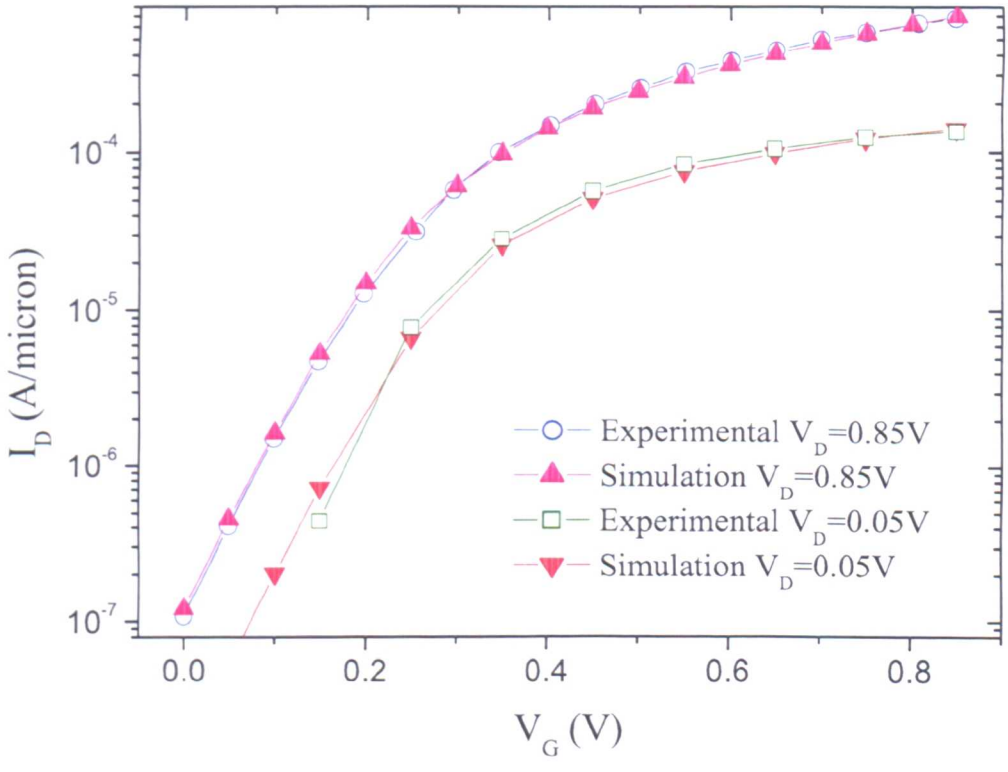


Figure 6.8: Calibrated I_D-V_G characteristics of the 35 nm physical gate length n-type Si MOSFET; experimental data from Inaba *et al.* [181]

The device is then simulated using Monte Carlo. The simulation commences with an assumption of a conventional Si channel with interface roughness parameters of Δ=0.5 nm and λ=1.8 nm, the results of which are plotted in diamond symbols in Fig. 6.9. The Taurus simulation results (with and without external resistances) are plotted in the same figure for comparison. Due to the unknown external resistance of the real device, it is expected that the Monte Carlo simulation should lie somewhere

[1] Taurus simulation results were provided by Mr. Fikru Adam-Lema, Device Modelling Group, University of Glasgow

Table 6.2: List of Monte Carlo simulated 35nm gate length MOSFETs

	Total effective strain	Process-induced	Using SiGe buffer	IR parameters
A	0%	0%	0%	Δ/λ–0.5/1.8
B	5%	5%	0%	Δ/λ–0.5/1.8
C	10%	5%	5%	Δ/λ–0.5/1.8
D	15%	5%	10%	Δ/λ–0.5/1.8
E	15%	5%	10%	Δ/λ–0.5/3.0
F	20%	5%	15%	Δ/λ–0.5/3.0

Note: the strain shown in percentage is equivalent to the strained by using SiGe buffer with Ge content in the same percentage

between the two Taurus curves. However, the Monte Carlo simulated drain currents (diamond symbols) are lower than the Taurus data, indicating a conventional Si channel underestimates the device performance.

However, it is known that process induced strain may occur in such small devices due to the use of contact-etch-stop cap layer, shallow trench isolation and silicidation processes [12,13,24]. We attempt here to simulate the effects of this to first order by assuming a strained Si body with different levels of strain. The simulated degrees of strain in the channel are assumed to be equivalent to that due to a SiGe buffer with 5%, 10% and 15% Ge content (represented by 5%, 10% and 15% Ge content equivalent strains). Table 6.2 lists the Monte Carlo simulated devices with different amount of strain. A conventional (unstrained) Si interface has been maintained for these devices (structures A-D in Table 6.2). The I_D-V_G curves for the devices with different amounts of strain are plotted in Fig. 6.9 at the same gate overdrive for a fair comparison. It is clear from the figure that the 35 nm Toshiba device is best modelled with 5% Ge content equivalent strain. The device performance as we could expect increases with increasing strain in the channel.

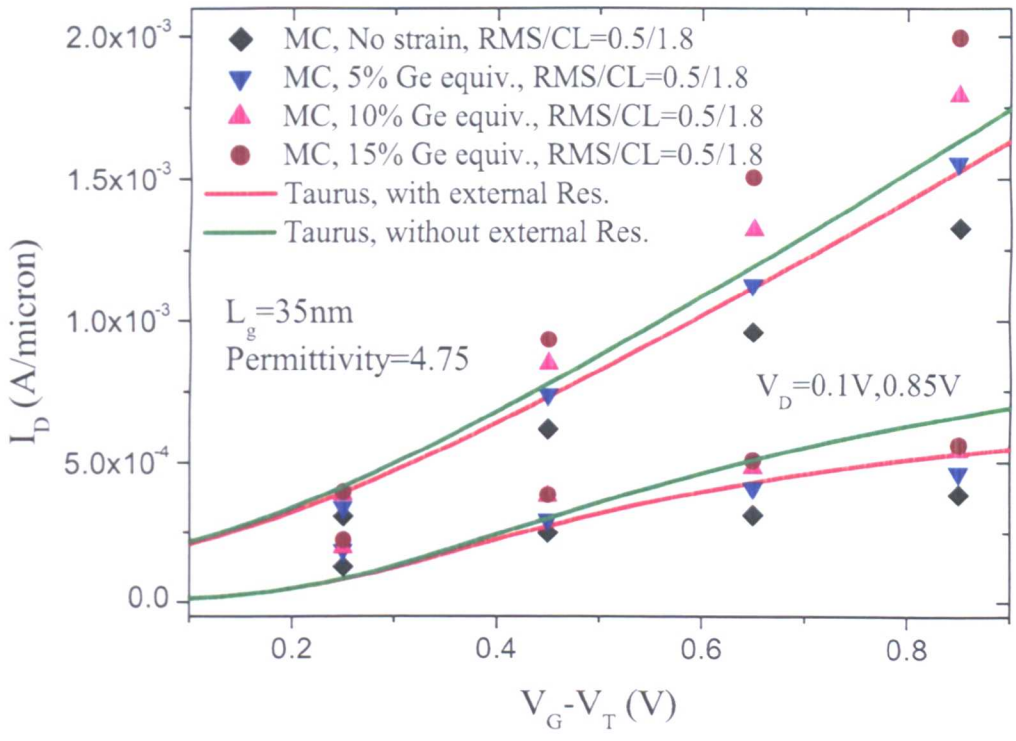


Figure 6.9: Monte Carlo calibrated I_D-V_G characteristics of 35 nm gate length n-type Si MOSFETs, compared to the data from calibrated Taurus simulator

To investigate the possible performance enhancement in scaled strained Si MOSFETs, Monte Carlo simulations are also carrier out for intentionally strained 35 nm devices (structures E and F in Table 6.2), assuming strained Si interface parameters (Δ=0.5 nm and λ=3.0 nm). Here, devices are assumed equivalent strains corresponding to substrate Ge contents of 15% and 20%. These equivalent strains are equal to the sum of the process induced (5% Ge content equivalent strain) and substrate strain, listed in Table 6.2. The device characteristics of all devices listed in Table 6.2 are shown in Fig. 6.10 for comparison. The drive current enhancement at V_G-V_T=0.85 V and V_D=0.85 V of the 35 nm strained Si channel MOSFET with 15% Ge content equivalent strain (process strained Si on a relaxed $Si_{0.9}Ge_{0.1}$ substrate) is about 32% over the 35 nm Toshiba device assuming only a process induced strain equivalent to a 5% Ge content, and 55% over the 35 nm Si MOSFET with zero strain in the channel. When considering higher degrees of strain in the channel,

for example, a strained Si channel with a total 20% Ge content equivalent strain, the 35 nm gate length strained Si MOSFET delivers a ~41% drive current enhancement over the Toshiba Si MOSFET, which is supported by the recently demonstrated 45% drive current enhancement in a 35 nm gate length strained $Si/Si_{0.8}Ge_{0.2}$ MOSFET [23].

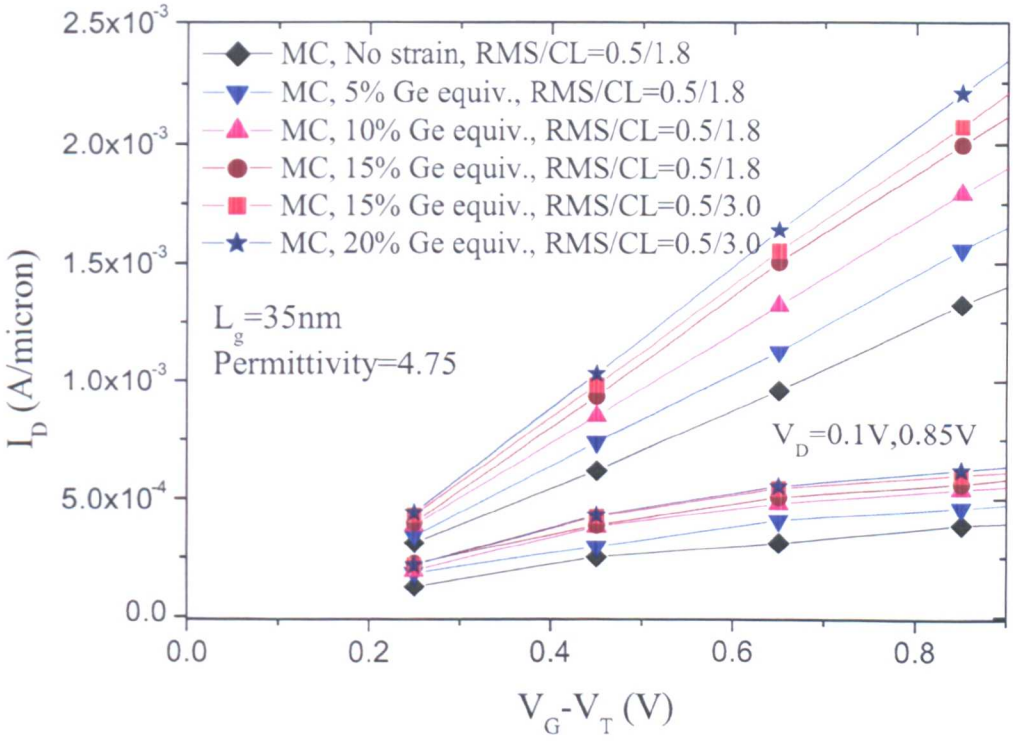


Figure 6.10: Monte Carlo simulated I_D-V_G characteristics of 35 nm gate length n-type Si MOSFETs with different strain within the channel

6.3 Performance Predictions of Strained Si MOSFETs with High-κ Dielectrics

6.3.1 Device Structure

The device structures simulated here are based on the 67 nm effective gate length conventional Si and strained Si MOSFETs calibrated in Chapter 5 (see Fig. 5.2)

and are assumed to have high-κ dielectrics with the same equivalent oxide thickness (EOT) of 2.2 nm. Fig. 6.11 illustrates the simulated device structures. Here, the two leading high-κ contenders, HfO_2 and Al_2O_3 [5], are studied. All other device parameters, such as the device geometry and the Si/insulator interface roughness parameters, are the same as the previous simulated SiO_2 devices, although this may be an optimistic scenario given the immaturity of the high-κ fabrication processes. The same EOT of different gate dielectrics leads to the same gate capacitances and enables identical electrostatic gate control between the devices with SiO_2 and high-κ dielectrics. However, in the presence of high-κ dielectrics carriers within the inversion layer are subjected to SO phonon scattering, leading to a reduction in the mobility of carriers. Therefore a reduction in device drive current is expected when high-κ dielectrics are introduced.

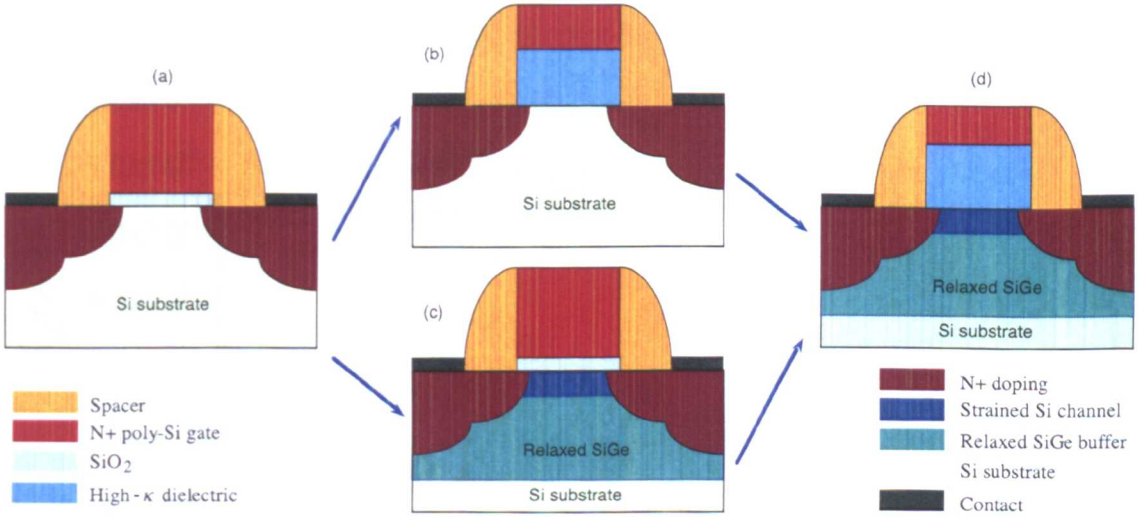


Figure 6.11: Schematics of the simulated devices: (a) conventional Si MOSFET with SiO_2; (b) conventional Si MOSFET with high-κ dielectrics; (c) strained Si MOSFET with SiO_2; (d) strained Si MOSFET with high-κ dielectrics

6.3.2 Performance Degradation due to Soft-optical Phonon Scattering

Soft optical phonon scattering has been included within the ensemble Monte Carlo simulator when studying the impact of the introduction of HfO_2 and Al_2O_3 in the gate stack. Fig. 6.12 illustrates the simulated I_D-V_G characteristics with and with-

out SO phonon scattering for the conventional Si MOSFET with a HfO_2 gate dielectric of 12.5 nm (EOT=2.2 nm). Fig. 6.13 illustrates the corresponding I_D-V_G characteristics of the strained Si MOSFET with an HfO_2 gate dielectric (EOT=2.2 nm). The simulations which include SO phonon scattering due to the HfO_2 gate dielectric exhibit a 40-50% reduction in the drive current at V_D=0.1 V and ~25% reduction at V_D=1.2 V. Similar percentage reductions are observed for both the Si and strained Si MOSFETs at the same gate overdrive, V_G-V_T=1.0 V. The reduction of the high-κ related current degradation at high V_D may be explained with reference to Fig. 6.2 which shows a reduction in the SO phonon scattering rate with the increasing carrier energy expected at high drain voltages.

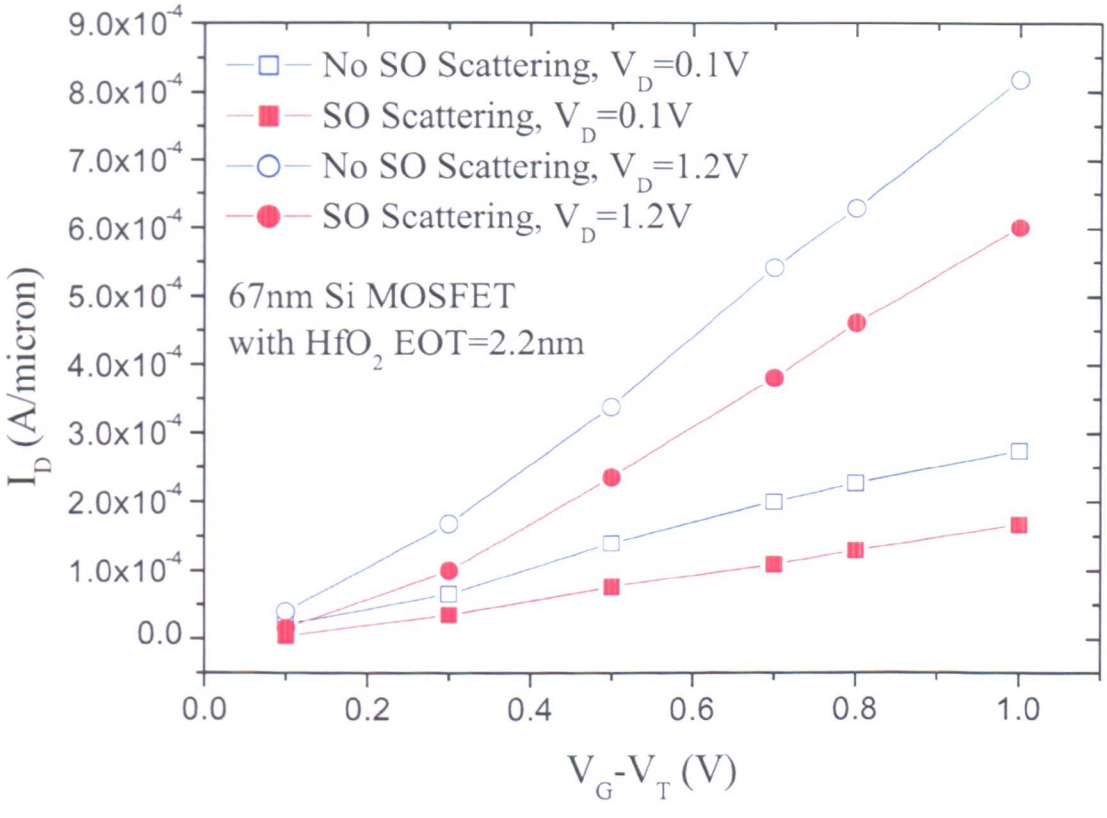


Figure 6.12: Monte Carlo simulated I_D-V_G characteristics with and without SO phonon scattering of the 67 nm n-type conventional Si MOSFET with HfO_2 gate dielectrics.

Fig. 6.14 shows the average channel velocities obtained from Monte Carlo simulations both with and without soft-optical phonon scattering at the same gate overdrive, V_G-V_T=1.0 V. SO phonon scattering significantly reduces the channel

velocities of both conventional and strained Si MOSFETs. However, it is observed that the channel velocity in the strained Si MOSFET with SO phonon scattering is slightly higher compared to the velocity in the conventional Si MOSFET without SO phonon scattering. This indicates that the introduction of high mobility strained channels could be used to counteract the performance degradation due to SO phonon scattering.

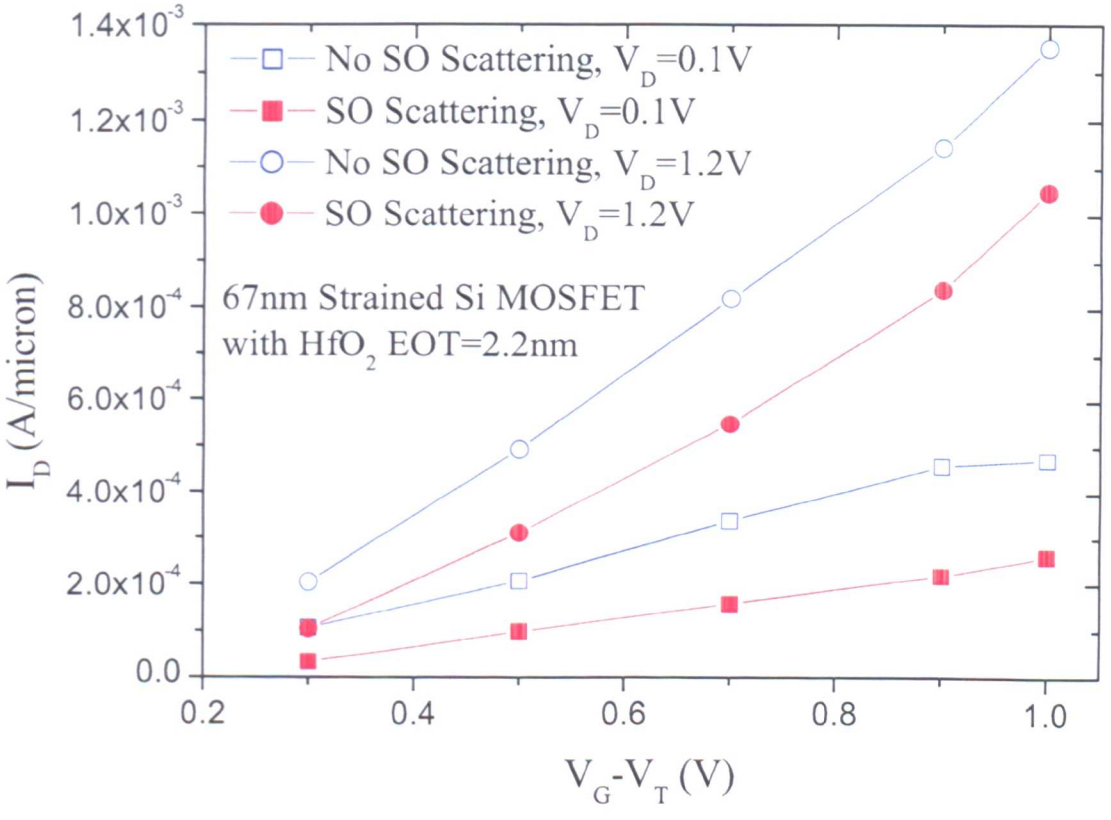


Figure 6.13: Monte Carlo simulated I_D-V_G characteristics with and without SO phonon scattering of the 67 nm strained Si MOSFET with HfO_2 gate dielectrics.

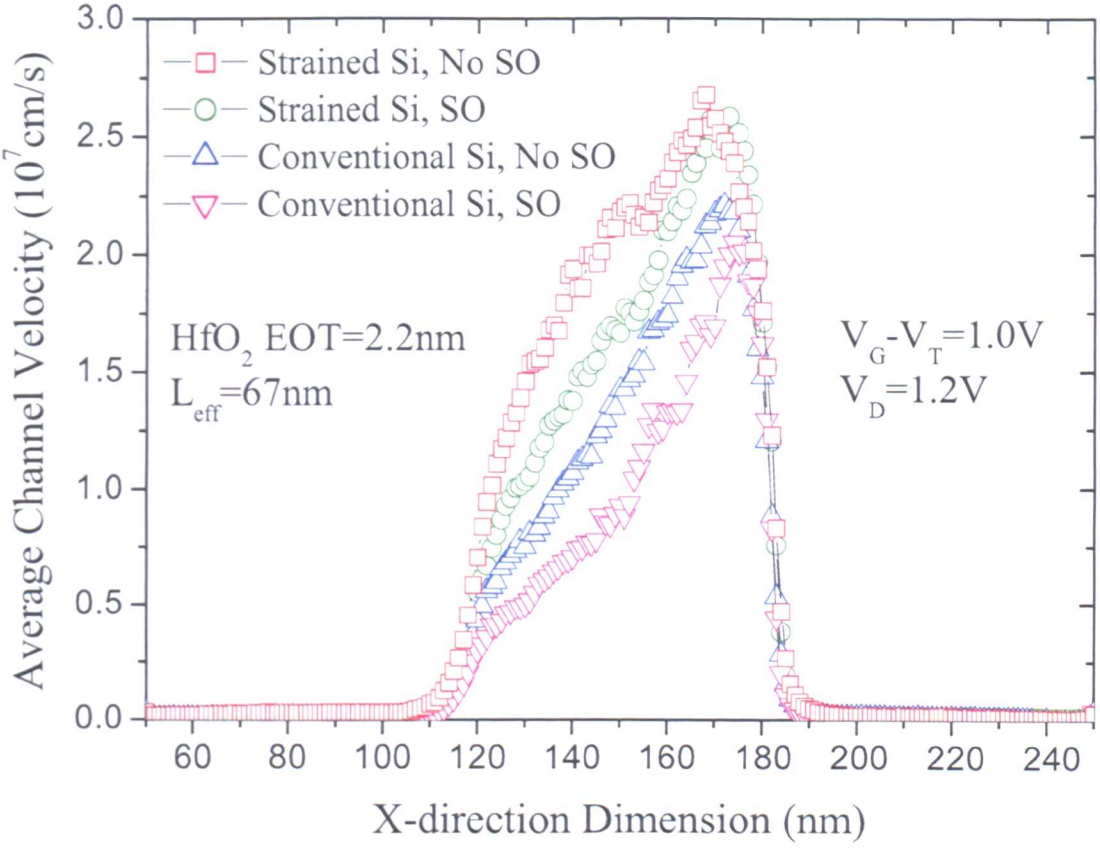


Figure 6.14: Average channel velocities obtained from MC simulations with and without SO phonon scattering in the 67 nm conventional and strained Si MOSFETs

Simulations were also performed to investigate the performance degradation associated with another promising high-κ candidate, Al_2O_3. Again comparisons were made between conventional Si and strained Si MOSFETs both with an Al_2O_3 gate stack of 7 nm (EOT=2.2 nm). The Monte Carlo simulated I_D-V_G characteristics with and without SO phonon scattering due to the Al_2O_3 dielectric at V_D=1.2 V are illustrated in Fig. 6.15. Compared to the 25% current degradation at V_G-V_T=1.0 V and V_D=1.2 V observed in the devices with HfO_2 gate dielectrics, it is observed in this case of the Al_2O_3 dielectric, a current reduction of around 10% for both conventional and strained Si MOSFETs. The differences in current degradation due to the SO phonon scattering can be explained by the larger phonon energies and a reduction in the difference between the static and optical permittivities (summarized in Table 6.1), in moving from HfO_2 to Al_2O_3 and SiO_2. In high-κ gate dielectrics, the

large static dielectric constant arises from the highly polarized ionic bonds, leading to lower phonon energies and smaller *optical permittivity*. The conventional SiO_2 has the lowest static dielectric constant, but harder bonds and thus higher phonon energies, which results in the small effect of SO phonon scattering in a Si/SiO_2 based device and is responsible for less than a 5% reduction in the drive current.

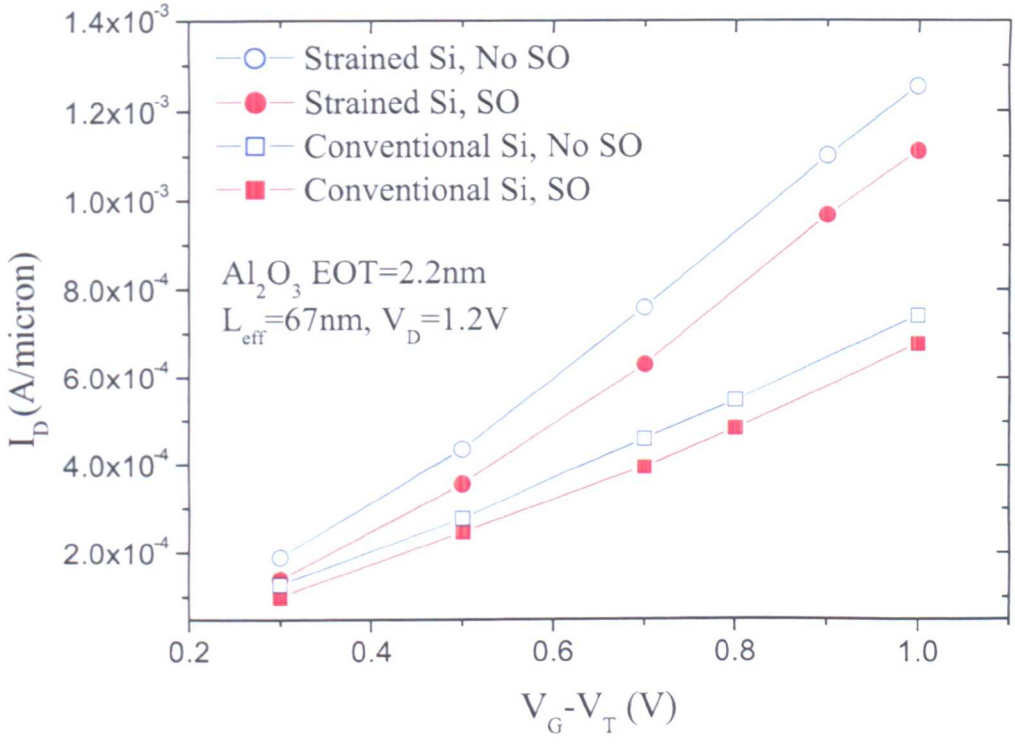


Figure 6.15: Performance comparisons of Monte Carlo simulations with and without soft optical phonon scattering in 67 nm n-type conventional and strained Si MOSFETs with Al_2O_3 gate dielectrics at V_D=1.2V

6.3.3 Other Mobility Degradation Sources

Aside from SO phonon scattering and geometrical interface roughness considered here, other scattering mechanisms present in the case of high-κ gate stacks may also contribute to the device performance degradation. Amongst them: remote Coulomb scattering due to interface states at the Si/insulator interface; fixed/trapped charges and interfacial dipoles, as well as the partial crystallization of the gate dielectric have

all been highlighted in the literature [7].

Another unknown factor is the quality (RMS height and correlation length) of the interface at the Si/insulator interface. The interface roughness parameters used in this study are assumed to be those of the Si/SiO_2 system. However, this may be too optimistic for the Si/high-κ interface given the relative immaturity of the high-κ growth compared to maturity and high quality of SiO_2 growth. Therefore, greater performance degradation than the predictions made in this work might be expected; the results presented here should be taken to represent the best case scenario.

However, the presence of a thin interfacial oxide between the high-κ gate dielectric and Si during growth may help to obtain a good quality interface and may partially suppress the effect of SO phonon scattering, although even a very thin layer would severly limit the equivalent oxide thickness of the high-κ gate stack. Recent experimental work of metal-gated devices with an HfO_2 gate stack has speculated that the screening by free carriers in the metal gate may help to suppress SO phonon scattering [96]. However, the degree of screening will depend on the gate material and the thickness and properties of the dielectrics. Further work is required to investigate this effect.

6.4 Summary

A non-perturbative semi-classical interface roughness scattering model incorporating roughness defined via an exponential autocorrelation function has been used to assess the effects of interface roughness in ultra small devices. The validation of the model is based on Monte Carlo simulations of the universal mobility behaviour of bulk Si and the device characteristics of sub-100 nm conventional Si IBM MOSFETs. It has been found that the Si/SiO_2 roughness for the conventional Si MOSFETs may be characterized by Δ=0.5 nm and λ=1.8 nm. Further calibrations of the model against the field-dependence of the electron mobility in strained Si and the device characteristics of the 67 nm strained Si IBM MOSFET indicate that reduced interface roughness scattering in the strained Si MOSFET is one of the key factors enhancing the device performance. This leads to the conclusion that strained Si may

have a smoother interface which is characterised by Δ–0.5 nm and λ–3.0 nm. The model was then used to study the 35 nm gate length Toshiba n-MOSFET where Monte Carlo simulation suggests that a 5% Ge content equivalent strain may occur within the channel. Using the calibrated roughness parameters for Si and strained Si, the 35 nm strained Si MOSFET with 15% and 20% Ge content equivalent strains deliver around 32% and 41% drive performance enhancement respectively over the 35 nm Toshiba Si MOSFET assuming a 5% Ge content equivalent strain induced by processing.

Monte Carlo simulations were also carried out to investigate the performance degradation due to soft-optical phonon scattering, which is becoming of great important with the introduction of high-κ gate dielectrics. Current degradation of around 25% at V_G-V_T–1.0 V and V_D–1.2 V is observed for the 67 nm effective gate length conventional and strained Si MOSFETs with a 2.2 nm EOT HfO_2 dielectric. This compares to a 10% degradation arising from the introduction of a 2.2 nm EOT Al_2O_3 dielectric. As a reference, the observed current degradation associated with SO phonons in SiO_2 in devices with identical structure is less than 5%. The results also indicate that the inherent mobility degradation associated with the high-κ gate stack MOSFETs could be compensated for by the introduction of strained Si channels. The infancy of high-κ gate fabrication techniques means that other performance degrading scattering mechanisms are likely to be present including a strong interface roughness contribution. Thus the overall performance degradation associated with high-κ gate dielectrics could well be worse than the predictions made in this chapter.

Chapter 7

Conclusions

Numerical simulations of conventional and strained Si MOSFETs were carried out in this work for RF and CMOS applications. This chapter summarizes the major contributions of this study and discusses future research directions.

7.1 Summary of Contributions

Different simulation techniques: Drift-Diffusion, Hydrodynamic and Monte Carlo simulations were used in this study for the simulations of strained Si/SiGe devices. In order to precisely describe the properties of strained Si and SiGe for the practical numerical simulation of related devices, parameters associated with the bandgap and band alignment of the Si/SiGe heterostructure have been calculated and summarized. The effect of strain on the band structure of Si have been analyzed and the strain induced band splitting of the degenerate subvalleys (thus reducing intervalley scattering) and the reduction in the effective mass are major reasons of the enhanced mobility in bulk strained Si. By comparing the existing theoretical and experimental data with this work, analytical expressions have been obtained for the bandgap; the band offset; the effective masses; the densities of states and the permittivity for the strained Si on relaxed $Si_{1-y}Ge_y$ and the strained $Si_{1-x}Ge_x$ on unstrained Si heterostructures.

Drift-Diffusion device simulations have been used to optimize the n-type buried strained Si channel Si/SiGe MODFET for RF and high linearity applications. Quali-

tative analysis and numerical simulations suggest that it is essential to consider both lateral and vertical device designs for RF performance and linearity. The gate-to-channel separation and gate to source/drain distances are found to have significant but opposite effects on device performance and linearity. The doping in the supply layers needs to be properly adjusted in order to obtain a high sheet carrier density in the channel and achieve better modulation efficiency. The doped channel device exhibits the best linearity but at the expense of reduced drive current, transconductance and RF performance. The simulations also show that scaling helps to improve RF performance, but slightly reduces device linearity. Therefore, trade-off designs are necessary for specific RF and/or high linearity applications.

Monte Carlo simulations show both electron and hole mobility enhancements in strained Si layers, indicating the advantage of using strained Si for CMOS applications. In the strained Si on relaxed SiGe hetero-system, when increasing the Ge content of the SiGe substrate, the electron and hole mobility enhancements saturate at Ge content of 15-20% and 35-40%, respectively. 1-D Poisson-Schrödinger solutions and Hydrodynamic device simulations show that in the sub-100 nm p-type strained Si (on relaxed SiGe buffer) MOSFET, a buffer Ge content between 30% and 40% is suitable for optimum device performance in conjunction with high doping ($>10^{17}$ cm^{-3}) in the channel region in order to suppress the parasitic conduction within the low-mobility SiGe layer.

Based on the successful calibrations in respect of the 80 nm gate length n-type and 90 nm gate length p-type conventional Si and strained Si MOSFETs fabricated by IBM, comprehensive Drift-Diffusion device simulations have been used to scale these devices down to 65 nm, 45 nm and 35 nm gate lengths and predict device and circuit behaviour of sub-100 nm strained Si MOSFETs for CMOS applications. The IBM MOSFET structure, with a single retrograde channel doping obtained from the calibrations against published data, shows a healthy performance improvement when the gate length is scaled down to 35 nm, minimising the short channel effects in the Si and strained Si transistors. Three-stage ring oscillators constructed with these scaled devices were simulated using the MEDICI device simulator. Strained Si circuits exhibit enhanced performance compared to their conventional Si counterparts. By

incorporating the SOI device structure, the simulations show that the strained Si SOI circuit delivers significant speed enhancement over all its Si competitors and exhibits reduced energy dissipation compared to its bulk counterparts. Since the Drift-Diffusion model used for the scaling cannot account for non-equilibrium transport which is likely to be present in small devices, the Hydrodynamic model and Monte Carlo were employed to predict the performance enhancement of the scaled strained Si devices. Compared to the 7-10% drive current enhancement observed in the original 90 nm p-type strained Si MOSFET (over its conventional Si counterpart), Hydrodynamic device simulations suggest that the current enhancement factor of the strained Si MOSFET (over corresponding conventional Si MOSFETs) increases by 10% when the gate length is scaled from 90 nm to 35 nm. Compared to the device structure with retrograde channel doping, the well-tempered strained Si MOSFETs with halo implants around the source/drain regions, which were simulated using the Hydrodynamic model, offers comparable drive current and well controlled short channel effects for effective gate lengths down to 25 nm.

Ensemble Monte Carlo simulations were used for predicting the device performance (drive current) of n-type strained Si MOSFETs. A non-perturbative interface roughness scattering model was used to investigate the impact of such scattering on the performance enhancement of strained Si MOSFETs. The model has been validated by calibrating with respect to the universal mobility behaviour and the experimental device characteristics of the 80 nm IBM conventional Si MOSFET. It is found that the Si/SiO_2 roughness for the conventional Si MOSFETs may be characterized by Δ–0.5 nm and λ–1.8 nm using this model. Further calibrations of the model against the field-dependence of the electron mobility in strained Si and the device characteristics of the 67 nm strained Si IBM MOSFET indicate that reduced interface roughness scattering in the strained Si MOSFET is one of the key factors enhancing the device performance. This leads to the conclusion that strained Si may have a smoother interface which is characterised by Δ–0.5 nm and λ–3.0 nm. The model was then used to study the 35 nm gate length n-type MOSFET fabricated by Toshiba. Monte Carlo simulation suggests that a 5% Ge content equivalent strain (see section 6.2.3) may occur within the channel. Using the calibrated roughness

parameters of Si and strained Si, a 35 nm strained Si MOSFET with 15% and 20% Ge content equivalent strains deliver about 32% and 41% drive performance enhancement respectively over the 35 nm Toshiba Si MOSFET which assumes a 5% Ge content equivalent strain induced by processing.

Monte Carlo simulations were also employed for investigating the performance degradation due to soft-optical phonon scattering, which arises with the introduction of high-κ gate dielectrics. Based on the calibrated 80 nm n-type conventional and strained Si IBM MOSFET structures, a device *current degradation* of around 25% at V_G-V_T=1.0 V and V_D=1.2 V is observed for both conventional and strained Si devices with a 2.2 nm EOT HfO_2 dielectric. This compares to an approximate 10% degradation arising from the introduction of a 2.2 nm EOT Al_2O_3 dielectric. As a reference, the current degradation associated with SO phonons in SiO_2 in devices with an otherwise identical structure is less than 5%. The results also indicate that the inherent mobility degradation associated with the high-κ gate stack MOSFET could be compensated for by the introduction of a strained Si channel. The infancy of high-κ gate fabrication techniques means that other performance degrading scattering mechanisms are likely to be present including a strong interface roughness contribution. Thus the overall performance degradation associated with high-κ gate dielectrics is expected to be worse than predicted in this study.

7.2 Future Work

The work on the Si/SiGe MODFETs for RF applications contained in this thesis is based on Drift-Diffusion device simulations, which cannot account for the self-heating effect due to low thermal conductivity of the thick SiGe buffer. The Hydrodynamic model may be used to simulate this effect along with velocity overshoot in the channel. Monte Carlo simulations may be more appropriate for these sub-100 nm devices in order to capture real carrier transport in the 2-DEG. The expected velocity overshoot in the strained Si channel should enhance RF performance. Replacing the few-microns thick SiGe buffer with a thin SiGe-on-insulator buffer, Si/SiGe MODFETs are expected to deliver enhanced performance, better

controlled short channel effects, and less leakage current. By replacing the thick SiGe buffer in the calibrated 70 nm n-type MODFET with a thin SiGe-on-insulator buffer, preliminary simulations show a 10% increase in f_T.

In this work, most of the simulations of strained Si MOSFETs for CMOS applications were based on the calibrated device structures with a single retrograde channel doping. For small devices, it is necessary to use complicated well and pocket profile designs. The scaling of n-type and p-type devices and the circuit simulations were based on Drift-Diffusion simulations. However, more advanced models such as the Hydrodynamic model and Monte Carlo simulations are more appropriate to predict the device performance of scaled devices. Based on the understanding of the strained Si MOSFETs from the numerical device simulations, compact modelling of strained Si CMOS is also required in order to properly assess the circuit behaviour and evaluate the advantages of strained Si CMOS.

For p-type MOSFETs, full band Monte Carlo simulations are required in order to precisely simulate hole transport based on the full description of the hole band structure, such as the six band $\boldsymbol{k \bullet p}$ theory, which has been employed in the Monte Carlo simulator available in the Device Modelling Group at the University of Glasgow. Further investigation of the interface roughness scattering model used in this study is also needed for p-type MOSFETs.

The simulated strained Si bulk MOSFET structure may also be integrated with new device structures such as SOI and double-gate MOSFETs. By incorporating strained Si within SOI or double-gate device structure, enhanced device performance is expected along with better controlled short channel effects and improved scalability. However, these changes of the device structure and material require more theoretical work in order to understand the carrier transport, such as phonon scattering, surface roughness scattering, remote Coulomb scattering, etc.

When MOSFETs are scaled beyond the 45 nm technology node, high-κ dielectrics become necessary. Apart from improving fabrication techniques to improve the quality of high-κ dielectrics, it is also necessary to investigate the new scattering mechanisms induced by the introduction of high-κ dielectrics, such as SO phonon scattering and remote Coulomb scattering. A preliminary simulation shows that the

drive current degradation is only around 8% in a 35 nm MOSFET with an HfO_2 dielectric of EOT−1nm, which is smaller than the performance degradation in the 80 nm Si and strained Si devices. This may be due to the SO phonon scattering rate drop with increasing energy. However, the unscreened SO phonon scattering model was used in this study. When the EOT oxide thickness becomes very small, a such scattering model is necessary that the screening effect from the gate with high concentration of free carriers can be accounted for.

Bibliography

[1] International Technology Roadmap for Semiconductors. Technical report, SEMATECH, 2003.

[2] B. Doris et al. Extreme scaling with ultra-thin Si channel MOSFETs. *Technical Digest of International Electron Devices Meeting*, pages 267 270, 2002.

[3] H. Wakabayashi et al. Sub-10-nm planer-bulk-CMOS devices using lateral juction control. *Technical Digest of International Electron Devices Meeting*, page 20.7, 2003.

[4] H. S. P. Wong. Beyond the conventional transistor. *IBM Journal of Research and Development*, 46(2/3):133 168, 2002.

[5] G. D. Wilk, R. M. Wallace, and J. M. Anthony. High-κ gate dielectrics: Current status and material properties considerations. *Journal of Applied Physics*, 89(10):5243 5275, 2001.

[6] M. V. Fischetti, D. A. Neumayer, and E. A. Cartier. Effective electron mobility in Si inversion layers in metal-oxide-semiconductor system with a high-κ insulator: The role of remote phonon scattering. *Journal of Applied Physics*, 90(9):4587 4608, 2001.

[7] Z. Ren, M. V. Fischetti, E. P. Gusev, E. A. Cartier, and M. Chudizk. Inversion channel mobility in high-κ high performance MOSFETs. *Technical Digest of International Electron Devices Meeting*, page 33.2, 2003.

[8] K. Kalna, L. Yang, and A. Asenov. Simulation study of III-V MOSFETs for digital applications. *Journal of Computational Electronics*, 2:341 346, 2003.

[9] E. H. C. Parker and T. E. Whall. SiGe heterostructure CMOS circuits and applications. *Solid State Electronics*, 43:1497 1506, 2000.

[10] T. E. Whall and E. H. C. Parker. SiGe - heterostructures for CMOS technology. *Thin Solid Films*, 367:250 259, 2000.

[11] http://www.intel.com/research/silicon/micron.htm.

[12] C.-H. Ge et al. Process-strained Si (PSS) CMOS technology feacturing 3D strain engineering. *Technical Digest of International Electron Devices Meeting*, page 3.7, 2003.

[13] V. Chan et al. High speed 45nm gate length CMOSFETs integrated into a 90nm bulk technology incorporating strain engineering. *Technical Digest of International Electron Devices Meeting*, page 3.8, 2003.

[14] R. People. Physics and application of Ge_xSi_{1-x}/Si strained layer heterostructure. *IEEE Journal of Quantum Electronics*, QE-22(9):1696 1710, 1986.

[15] J. Welser, J. L. Hoyt, and J. F. Gibbons. NMOS and PMOS transistors fabricated in strained silicon/relaxed silicon-germanium structures. *Technical Digest of International Electron Devices Meeting*, page 31.7, 1992.

[16] http://www.research.ibm.com/resources/press/strainedsilicon/.

[17] J. R. Huang et al. Performance of 70nm strained-silicon CMOS devices. *Symposium on VLSI Technology Digest of Technical Papers*, pages 103 104, 2003.

[18] S. Takagi, T. Mizuno, and N. Sugiyama. Strained-Si-on-illustrator (strained-SOI) MOSFETs - Concepts, device structure and device characteristics. *IEICE Transaction on Electronics*, EC84-C(8):1043 1049, 2001.

[19] T. A. Langdo, A. Locktefeld, and M. T. Curie. Preparations of novel SiGe-free strained Si on illustrator substrates. *Proceeding of IEEE International SOI Conference*, pages 211 212, 2002.

[20] K. Rim et al. Fabrications and mobility characteristics of ultra-thin strained Si directly on illustrator (SSDOI) MOSFETs. *Technical Digest of International Electron Devices Meeting*, page 3.1, 2003.

[21] K. Rim et al. Characteristics and device design of sub-100nm strained Si N- and P-MOSFETs. *Symposium on VLSI Technology Digest of Technical Papers*, pages 98 101, 2002.

[22] M. Yang et al. High performance CMOS fabricated on hybrid substrate with different crystal orientations. *Technical Digest of International Electron Devices Meeting*, page 18.7, 2003.

[23] J.-S. Goo et al. Scalability of strained-Si nMOSFETs down to 25nm gate length. *IEEE Electron Device Letters*, 24(5):351 353, 2003.

[24] H. C. H. Wang et al. Substrate-strained silicon technology - process integration. *Technical Digest of International Electron Devices Meeting*, page 3.4, 2003.

[25] M. Enciso-Aguilar et al. Dc and high frequency performance of $0.1\mu m$ n-type $Si/Si_{0.6}Ge_{0.4}$MODFET with f_{max}—188GHz at 300K and f_{max}—230GHz at 50K. *Electronics Letters*, 39(1):149 151, 2003.

[26] M. Enciso-Aguilar et al. 0.3dB minimum noise figure at 25GHz of $0.13\mu m$ $Si/Si_{0.58}Ge_{0.42}$ n-MODFET. *Electronics Letters*, 37(17):1089 1090, 2001.

[27] http://www.synopsys.com.

[28] G. Snider. http://www.nd.edu/~gsnider/.

[29] J. R. Watling, Y. P. Zhao, A. Asenov, and J. Barker. Non-equilibrium hole transport in deep sub-micron well-tempered Si p-MOSFETs. *VLSI Design*, 13:169 173, 2001.

[30] L. Yang et al. Scaling study of Si/SiGe MODFETs for RF applications. *Proceeding of the 10th IEEE International Symposium on Electron Devices for Microwave and Optoelectronic Applications*, pages 101 106, 2002.

[31] K. Rim et al. Strained Si NMOSFETs for high performance CMOS technology. *Symposium on VLSI Technology Digest of Technical Papers*, pages 59 62, 2001.

[32] L. Yang et al. A study of high linearity Si/SiGe HFETs. *Proceedings of the 14th Workshop on Modelling and Simulation of Electron Device*, pages 41 44, 2003.

[33] L. Yang et al. Optimizations of sub-100nm strained Si/SiGe MODFETs for high linearity RF applictions. *Proceedings of the 2003 IEEE Conference on Electron Devices and Solid-State Electronics*, pages 331 334, 2003.

[34] L. Yang et al. Simulations of scaled sub-100nm strained Si/SG e p-channel MOSFETs. *Journal of Computational Electronics*, 2:363 368, 2003.

[35] L. Yang et al. Reduced interface roughness in sub-100nm strained Si n-MOSFETs - A Monte Carlo simulation study. *Proceeding of the 5th European Workshop on Ultimate Integration of Silicon*, pages 23 26, 2004.

[36] M. Boriçi et al. A non perturbative model of surface roughness scattering for Monte Carlo simulation of relaxed silicon n-MOSFETs. *Journal of Computational Electronics*, 2:163 168, 2003.

[37] M. Johansson et al. HfO_2 gate dielectrics on strained-Si and strained-SiGe layers. *Semiconductor Science and Technology*, 18:820 826, 2003.

[38] Z. Shi, D. Onsongo, K. Onishi, J. C. Lee, and S. K.Banerje. Mobility enhancement in surface channel SiGe PMOSFETs with HfO_2 gate dielectrics. *IEEE Electron Device Letters*, 24:34 36, 2003.

[39] D. Wu et al. A novel strained $Si_{0.7}Ge_{0.3}$ surface-channel pMOSFET with an ALD $TiN/Al_2O_3/HfAlO_xAl_2O_3$ gate stack. *IEEE Electron Device Letters*, 24:171 173, 2003.

[40] T. Tezuka, N. Sugiyama, T. Mizuno, and S. Takagi. Ultrathin body SiGe-on-insulator pMOSFETs with high-mobility SiGe surface channels. *IEEE Transactions on Electron Devices*, 50:1328 1335, 2003.

[41] Y. Yeo et al. SiGe channel p-MOSFETs scaling-down. *IEEE Transactions on Electron Devices*, 49:279 287, 2002.

[42] F. Andrieu et al. Design and fabrication of 50nm thin-body p-MOSFETs with a SiGe heterostructure. *Proceeding of ESSDERC*, pages 267 270, 2003.

[43] J. R. Watling, L. Yang, A. Asenov, J. R. Barker, and S. Roy. The impact of interface roughness and degeneracy in relaxed and strained Si n-MOSFETs. *Solid States Electronics*, 48:1337 1346, 2004.

[44] F. M. Bufler, P. Grak, S. Keith, and B. Meinerzhagen. Full band Monte Carlo investigation of electron transport in strained Si grown on $Si_{1-x}Ge_x$ substrates. *Applied Physics Letters*, 71:3477 3499, 1997.

[45] T. Mizuno et al. (110)-Surface Strained-SOI CMOS Devices with Higher Carrier Mobility. *Symposium on VLSI Technology Digest of Technical Papers*, pages 97 98, 2003.

[46] U. König and H. Dämbkes. SiGe HBTs and HFETs. *Solid-State Electronics*, 38(9):1595 1602, 1995.

[47] K. L. Wang, S. G. Thomas, and M. O. Tanner. SiGe band engineering for MOS, CMOS and quantum effect devices. *Journal of Materials Science: Materials in Electronics*, 6:311 324, 1995.

[48] F. Schäffler. High-mobility Si and Ge structures. *Semiconductor Science Technology*, 12:1515 1549, 1997.

[49] T. E. Whall and E. H. C. Parker. SiGe heterostructure for FET applications. *Journal of Physics D-Applied Physics*, 31(12):1397 1416, 1998.

[50] D. J. Paul. Silicon Germanium heterostructures in electronics: the present and the future. *Thin Solid Films*, 321:172 180, 1998.

[51] U. König, A. J. Boers, F. Schaäffler, and E Kasper. Enhancement mode n-Channel Si/SiGe MODFET with high intrinsic transconductance. *Electronics Letters*, 28(2):160 162, 1992.

[52] U. König and A. Gruhle. High frequency SiGe heterostructure devices. *Proceeding of IEEE Cornell Conference of Advanced Concepts in High Speed Semiconductor Devices and Circuits*, pages 14 23, 1997.

[53] U. König et al. Design rules for n-type SiGe hetero FETs. *Solid State Electronics*, 41(10):1541 1547, 1997.

[54] F. Aniel et al. High performance 100nm T-gate strained $Si/Si_{0.6}Ge_{0.4}$ n-MODFET. *Solid State Electronics*, 47:283 289, 2003.

[55] Q. Ouyang et al. A comprehensive simulation study of strained-Si/SiGe nMODFET scaling for RF applications. *Proceeding of SISPAD*, pages 59 62, 2002.

[56] E. A. Fitzgerald et al. Totally relaxed Ge_xSi_{1-x} layers with low threading dislocation densities grown on Si substrates. *Appllied Physics Letter*, 59(7):811 813, 1991.

[57] G. Kissinger, T. Morgenstem, G. Morgenstem, and H. Richter. Stepwise equilibrated graded Ge_xSi_{1-x} buffer with very low threading dislocation density on Si (001). *Appllied Physics Letter*, 66(16):2083 2085, 1995.

[58] T. Hackbarth, H. Kibbel, M. Glueck, G. Hoeck, and H. J. Herzog. Artificial sbstrates for n- and p-type SiGe heterostructre field-effect transistors. *Thin Solid Films*, 321:136 140, 1998.

[59] T. Hackbarth, G. Hoeck, H. J. Herzog, and M. Zeuner. Strain relieved SiGe buffers for Si-based heterostructure field-effect transistors. *Journal of Crystal Growth*, 201/202:734 738, 1999.

[60] H. J. Herzog, T. Hackbarth, G. Höck, M. Zeuner, and U. König. SiGe-based FETs: buffer issues and device results. *Thin Solid Films*, 380:36 41, 2000.

[61] K. A. Jenkins and K. Rim. Measurement of the effect of self-heating in strained-Silicon MOSFETs. *IEEE Electron Device Letters*, 23(6):360 362, 2002.

[62] T. Hackbarth et al. Alternatives to thick MBE-grown relaxed SiGe buffers. *Thin Solid Films*, 369:148 151, 2000.

[63] A. R. Powell, S. S. Iyer, and F. K. LeGoues. New approach to the growth of low dislocation relaxed SiGe material. *Appllied Physics Letter*, 64(14):1856 1858, 1994.

[64] F. K. LeGoues, A. Powell, and S. S. Iyer. Relaxation of SiGe thin films grown on Si/SiO_2 substrates. *Journal of Appllied Physics*, 75(11):7240 7242, 1994.

[65] A. G. O'Neill and D. A. Antoniadis. Investigation of Si/SiGe-based FET geometries for high frequency performance by computer simulation. *IEEE Transactions on Electron Devices*, 44(1):80 88, 1997.

[66] W. Lu et al. High performance 0.1μm gate-length p-Type SiGe MODFET's and MOS-MODFET. *IEEE Transactions on Electron Devices*, 47(8):1645 1652, 2000.

[67] J. Welser, J. L. Hoyt, S. Takagi, and J. F. Gibbons. Strain dependence of the performance enhancement in strained Si n-MOSFETs. *Technical Digest of International Electron Devices Meeting*, page 15.2, 1994.

[68] K. Rim, J. Welser, J. L. Hoyt, and J. F. Gibbons. Enhanced hole mobilities in strained Si surface channel p-MOSFETs. *Technical Digest of International Electron Devices Meeting*, page 20.3, 1995.

[69] K. Rim, J. L. Hoyt, and J. F. Gibbons. Transconductance enhancement in deep submicron strained-Si n-MOSFETs. *Technical Digest of International Electron Devices Meeting*, page 26.8, 1998.

[70] M. V. Fischetti, F. Gámiz, and W. Hänsch. On the enhanced electron mobility in strained-silicon inversion layers. *Journal of Applied Physics*, 92(12):7320 7324, 2002.

[71] K. Rim et al. Strained Si CMOS (SS CMOS) technology: opportunities and challenges. *Proceedings of the 3rd European Workshop on Ultimate Integration of Silicon*, pages 73 76, 2002.

[72] J. L. Hoyt et al. Strained silicon mosfet technology. *Technical Digest of International Electron Devices Meeting*, pages 23 26, 2002.

[73] N. Braga, A. Buczkowski, H. R. Kirk, and G. A. Rozgonyi. Formation of cylindrical n/p junction diodes by arsenic enhanced diffusion along interfacial misfit dislocations in p-type epitaxial Si/Si(Ge). *Applied Physics Letters*, 64(11):1410 1412, 1994.

[74] S. Eguchi, C. W. Leitz, E. A. Fitzgerald, and J. L. Hoyt. Diffusion behavior of ion-implanted n-type dopants in silicon germanium. *Materials Research Society Symposium Proceedings*, 686:33 38, 2002.

[75] P. Kuo, J. L. Hoyt, J. F. Gibbons, J. E. Turner, and D. Lefforge. Boron diffusion in Si and $Si_{1-x}Ge_x$. *Materials Research Society Symposium Proceedings*, 379:373 378, 1995.

[76] 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. *Applied Physics Letters*, 79(14):2148 2150, 2001.

[77] J. G. Fiorenza et al. Film thickness constraints for manufacturable strained silicon CMOS. *Semiconductor Science and Technology*, 19:L4 L8, 2004.

[78] S. Takagi et al. Channel structure design, fabrication and carrier transport properties of strained-Si/SiGe-on-illustrator (strained SOI MOSFETs. *Technical Digest of International Electron Devices Meeting*, page 3.3, 2003.

[79] H. Yin et al. Fully depleted strained-Si on illustrator NMOSFETs without relaxed SiGe buffers. *Technical Digest of International Electron Devices Meeting*, page 3.2, 2003.

[80] S. Fukatsu, Y. Ishikawa, T. Saito, and N. Shibata. SiGe-based semiconductor-on-insulator substrate created by low-energy separation-by-implanted-oxygen. *Applied Physics Letters*, 72(26):3485 3487, 1998.

[81] Y. Ishikawa, N. Shibata, and S. Fukatsu. SiGe-on-insulator substrate using SiGe alloy grown Si(001). *Applied Physics Letters*, 75(7):983 985, 1999.

[82] N. Sugiyama et al. Formation of strained-Silicon layer on thin relaxed-SiGe/SiO_2/Si structure using SIMOX technology. *Thin Solid Films*, 369:199 202, 2000.

[83] L.-J. Huang et al. Carrier mobility enhancement in strained Si-on-insulator fabricated by wafer bonding. *Symposium on VLSI Technology Digest of Technical Papers*, pages 57 58, 2001.

[84] T. Tezuka, N. Sugiyama, , and S. Takagi. Fabrication of strained Si on an ultrathin SiGe-on- insulator virtual substrate with a high-Ge fraction. *Applied Physics Letters*, 79(12):1798 1800, 2001.

[85] T. Tezuka, N. Sugiyama, T. Mizuno, and S. Takagi. Novel fully depleted SiGe-on-illustrator pMOSFETs with high mobility SiGe channels. *Technical Digest of International Electron Devices Meeting*, page 33.6, 2001.

[86] T. Tezuka, N. Sugiyama, T. Mizuno, and S. Takagi. High-performance Strained Si-on-insulator MOSFETs by novel fabrication processes utilizing Ge-condensation technique. *Symposium on VLSI Technology Digest of Technical Papers*, pages 96 97, 2002.

[87] T. Mizuno et al. High performance CMOS operation of strained-SOI MOSFETs using thin film SiGe-on-insulator substrate. *Symposium on VLSI Technology Digest of Technical Papers*, pages 106 107, 2002.

[88] T. Mizuno et al. High performance strained-SOI CMOS devices using thin film SiGe-on-insulator technology. *IEEE Transactions on Electron Devices*, 50(4):988 994, 2003.

[89] S. Takagi et al. Device structure and electrical characteristics of strained-Si-on-illustrator (strained-SOI) MOSFETs. *Material Science and Engineering*, B89:426 434, 2002.

[90] T. Mizuno, N. Sugiyama, T. Satake, and S. Takagi. Advanced SOI-MOSFETs with strained Si channel for high speed CMOS - electron/hole mobility enhancement. *Symposium on VLSI Technology Digest of Technical Papers*, pages 210 211, 2000.

[91] B. H. Lee et al. Performance enhancement on sub-70nm strained silicon SOI MOSFETs on ultra-thin thermally mixed strained Si/SiGe on illustraotr (TM-SOI) substrate with raised S/D. *Technical Digest of International Electron Devices Meeting*, page 17.7, 2002.

[92] L.-J. Huang et al. Electron and hole mobility enhancement in strained SOI by wafer bonding. *IEEE Transactions on Electron Devices*, 49(9):1566 1571, 2002.

[93] T. A. Langdo et al. SiGe-free strained Si on insulator by wafer bonding and layer transfer. *Applied Physics Letters*, 82(24):4256 4258, 2003.

[94] I. Lauer et al. Fully depleted n-MOSFETs on supercritical thickness strained SOI. *IEEE Electron Device Letters*, 25(2):83 85, 2004.

[95] K. Rim et al. Mobility enhancement in strained Si NMOSFETs with HfO_2 gate dielectrics. *Symposium on VLSI Technology Digest of Technical Papers*, pages 12 13, 2002.

[96] S. Datta et al. High mobility Si/SiGe strained channel MOS transistors with HfO_2/TiN gate stack. *Technical Digest of International Electron Devices Meeting*, page 28.1, 2003.

[97] W. Zhu, J.-P. Han, and T. P. Ma. Mobility measurement and degradation mechanisms of MOSFETs made with ultrathin high-κ dielectrics. *IEEE Transactions on Electron Devices*, 51(1):98 105, 2004.

[98] W. Zhu and T. P. Ma. Temperature dependence of channel mobility in HfO_2-gated NMOSFETs. *IEEE Electron Device Letters*, 25(2):89 91, 2004.

[99] F. Gámiz, J. B. Roldán, and J. A. López-Villanueva. Phonon-limited electron mobility in ultrathin silicon-on-insulator inversion layers. *Journal of Applied Physics*, 83(9):4802–4806, 1998.

[100] Masanari Shoji and Seiji Horiguchi. Electronic structures and phonon-limited electron mobility of double-gate silicon-on-insulator Si inversion layers. *Journal of Applied Physics*, 85(5):2722–2731, 1999.

[101] M. V. Fischetti and S. E. Laux. Long-range Coulomb interactions in small Si devices. Part I: Performance and reliability. *Journal of Applied Physics*, 89(2):1205–1230, 2001.

[102] M. V. Fischetti. Long-range Coulomb interactions in small Si devices. Part II: Effective electron mobility in thin-oxide structures. *Journal of Applied Physics*, 89(2):1232–1250, 2001.

[103] S. Saito, K. Torii, M. Hiratani, and Takahiro Onai. Improved theory for remote-charge-scattering-limited mobility in metal-oxide-semiconductor transistors. *Applied Physics Letters*, 81(13):2391–2393, 2002.

[104] F. Gámiz, J. B. Roldán, J. E. Carceller, and P. Cartujo. Monte Carlo simulation of remote-Coulomb-scattering-limited mobility in metal-oxide-semiconductor transistors. *Applied Physics Letters*, 82(19):2351–2353, 2003.

[105] D. Esseni and A. Abramo. Modeling of electron mobility degradation by remote Coulomb scattering in ultrathin oxide MOSFETs. *IEEE Transactions on Electron Devices*, 50(7):1665–1674, 2003.

[106] F. Gámiz, J. B. Roldán, J. A. López-Villanueva, et al. Surface roughness at the Si-SiO_2 interfaces in fully depleted silicon-on-insulator inversion layers. *Journal of Applied Physics*, 86(12):6854–6859, 1999.

[107] D. Esseni, A. Abramo, L. Selmi, and E. Sangiorgi. Study of low field electron transport in ultra-thin single and double gate SOI MOSFETs. *Technical Digest of International Electron Devices Meeting*, pages 719–722, 2002.

[108] Masanari Shoji and Seiji Horiguchi. Phonon-limited inversion layer electron mobility in extremely thin Si layer of silicon-on-insulator metal-oxide-semiconductor field-effect transistor. *Journal of Applied Physics*, 82(12):6096 6101, 1997.

[109] F. Gámiz and M. V. Fischetti. Monte Carlo simulation of double-gate silicon-on-insulator inversion layers: The role of volume inversion. *Journal of Applied Physics*, 89(10):5478 5487, 2001.

[110] M. T. Currie et al. Carrier mobilities and process stability of strained Si n- and p-MOSFETs on SiGe virtual substrates. *Journal of Vacuum Science and Technology B*, 19(6):2268 2279, 2001.

[111] F. Gámiz, J. B. Roldán, and A. Godoy. Strained-Si/strained-Si-on-illustrator inversion layers: the role of strained-Si layer thickness on electron mobility. *Applied Physics letters*, 80(22):4160 4162, 2002.

[112] F. Gámiz, P. Cartujo-Cassinello, J. B. Roldán, and F. Jieénez-Molinos. Electron transport in strained Si inversion layers grown on SiGe-on-insulator substrates. *Journal of Applied Physics*, 92(1):288 295, 2002.

[113] C. W. Leitz et al. Hole mobility enhancements in strained Si/$Si_{1-y}Ge_y$ p-type metal-oxide-semiconductor field-effect transistors grown on relaxed $Si_{1-y}Ge_y$ (x<y) virtual substrates. *Applied Physics Letters*, 79(25):4246 4248, 2001.

[114] M. L. Leea and E. A. Fitzgerald. Optimized strained Si/strained Ge dual-channel heterostructures for high mobility p- and n-MOSFETs. *Technical Digest of International Electron Devices Meeting*, page 18.1, 2003.

[115] C. O. Chui, H. Kim, P. C. McIntyre, and K. C. Saraswat. A germanium NMOSFET process integrating metal gate and improved high-κ dielectrics. *Technical Digest of International Electron Devices Meeting*, page 18.3, 2003.

[116] A. Ritenour et al. Epitaxial strained germanium p-MOSFETs with HfO_2 gate dielectric and TaN gate electrode. *Technical Digest of International Electron Devices Meeting*, page 18.2, 2003.

[117] http://www.silvaco.com.

[118] http://www.ise.com.

[119] R. Braunstein, A. R. Moore, and F. Herman. Intrinsic optical absorption in germanium-silicon alloys. *Physical Review*, 109(3):695 710, 1958.

[120] C. G. Van de Walle and R. M. Martin. Theoretical calculations of heterojunction discontinuities in the Si/Ge system. *Physical Review B*, 34(8):5621 5634, 1986.

[121] M. M. Rieger and P. Vogl. Electronic-band parameters in strained $Si_{1-x}Ge_x$ alloys on $Si_{1-y}Ge_y$ substrates. *Physical Review B*, 48(19):14276 14287, 1993.

[122] E. Kasper and K. Lyutovich, editors. *Properties of Silicon Germanium and SiGe:Carbon.* IEE INSPEC, 2000.

[123] U. Ravaioli. Hierarchy of simulation approaches for hot carrier transport in deep submicron devices. *Semiconductor Science and Technology*, 13:1 10, 1998.

[124] S. Selberherr. *Analysis and Simulation of Semiconductor Devices.* Springer-Verlag, 1984.

[125] Y. Zhao. *Simulation and optimisation of SiGe MOSFETs.* PhD thesis, University of Glasgow, 2000.

[126] C. G. Van de Walle and R. M. Martin. Theoretical calculations of semiconductor heterojunction discontinuities. *Journal of Vaccum Science Technology B*, 4(4):1055 1059, 1986.

[127] L. Colombo, R. Resta, and S. Baroni. Valence-band offsets at strained Si/Ge interfaces. *Physical Review B*, 44(1):5572 5579, 1991.

[128] G. P. Schwartz et al. Core-level photoemission measurements of valence-band offsets in highly strained heterojuntions: Si-Ge systems. *Physical Review B*, 39(2):1235 1241, 1989.

[129] E. T. Yu et al. Measurement of the valence band offset in novel heterojunction systems: Si/Ge(100) and AlSb/ZnTe(100). *Journal of Vaccum Science Technology B*, 8(4):908 915, 1990.

[130] J. F. Morar, P. E. Batson, and J. Tersoff. Heterojunction band lineups in Si-Ge alloys using spatially resolved electron-energy-loss spectroscopy. *Physical Review B*, 47(7):4107 4110, 1993.

[131] W.-X. Ni and G. V. Hansson. Band offsets in pseudomorphically grown si $Si_{1-x}Ge_x$ heterostructures studied with core-level x-ray photoelectron spectroscopy. *Physical Review B*, 42(5):3030 3037, 1990.

[132] S. Takagi et al. Evaluation of the valence band discontinuity of $Si/Si_{1-x}Ge_x/Si$ heterostructures by application of admittance spectroscopy to MOS capacitors. *IEEE Transactions on Electron Devices*, 45(2):494 501, 1998.

[133] J. Weber and M. I. Alonso. Near-band-gap photoluminescence of Si-Ge alloys. *Physical Review B*, 40(8):5683 5693, 1989.

[134] S. Krishnamurthy and A. Sher. Band structures of $Si_{1-x}Ge_x$ alloys. *Physical Review B*, 33(2):1026 1035, 1986.

[135] S. M. Sze. *Semiconductor Devices: Physics and Technology*. Wiley, 1985.

[136] D. J. Robbins, L. T. Canham, S. J. Barnett, A. D. Pitt, and P. Calcott. Near-band-gap photoluminescence from pseudomorphic $Si_{1-x}Ge_x$ single layers on silicon. *Journal of Applied Physics*, 71(3):1407 1414, 1992.

[137] D. Dutartre, G. Brämond, A. Souifi, and T. Benyattou. Excitonic photoluminescence from Si-capped strained $Si_{1-x}Ge_x$ layers. *Physical Review B*, 44(20):11525 11527, 1991.

[138] J. C. Bean. Silicon-based semiconductor heterostructures: Column IV bandgap engineering. *Proceedings of the IEEE*, 80(4):571 587, 1992.

[139] J. Eberhardt and E. Kasper. Bandgap narrowing in strained SiGe on the basis of electrical measurements on Si/SiGe/Si hetero bipolar transistors. *Material Science and Engineering*, B89:93–96, 2002.

[140] D.V. Lang, R. People, J. C. Bean, and A. M. Sergent. Measurement of the band gap of Ge_xSi_{1-x}/Si strained-layer heterostructure. *Applied Physics Letters*, 47(12):1333–1335, 1985.

[141] J. Spitzer, K. Thonke, R. Sauer, et al. Direct observation of band-edge luminescence and alloy luminescence from ultrametastable silicon-germanium alloy layers. *Applied Physics Letters*, 60(14):1729–1731, 1992.

[142] M. V. Fischetti and S. E. Laux. Band structure, deformation potentials, and carrier mobility in strained Si, Ge and SiGe alloys. *Journal of Applied Physics*, 80(4):2234–2252, 1996.

[143] Synopsys Inc. *MEDICI Manual 2003.6*, 2003.

[144] J. W. Harrison and J. R. Hauser. Theoretical calculations of electron mobility in ternary III-V compounds. *Journal of Applied Physics*, 47:292–300, 1976.

[145] Synopsys Inc. *TAURUS Manual 2003.4*, 2003.

[146] S. E. Laux. Techniques for small-signal analysis of semiconductor devices. *IEEE Transactions on Electron Devices*, 32(10):2028–2037, 1985.

[147] Roer and Theo G. van de. *Microwave Electronic Devices*. Chapman and Hall Press, 1994.

[148] S. Babiker et al. Complete Monte Carlo RF analysis of real short-channel compound FET's. *IEEE Transactions on Electron Devices*, 45(8):1644–1652, 1998.

[149] M. Berroth and R. Bosch. Broad-band determination of the FET small-signal equivalent circuit. *IEEE Transactions on Microwave Theory and Techniques*, 38(7):891–895, 1990.

[150] M. Zeuner, T. Hackbarth, *G. Höck, D. Behammer, and U. König.* High-frequency SiGe-n-MODFET for microwave applications. *IEEE Microwave and Guided Wave Letters*, 9(10):410 412, 1999.

[151] M. S. Gupta. Power gain in feedback amplifiers: A classic revisited. *IEEE Transactions on Microwave Theory and Techniques*, 40(5):864 879, 1992.

[152] B. Razavi. *RF Microelectronics.* Prentice-Hall, 1998.

[153] K. Rim, J. L. Hoyt, and J. F. Gibbons. Fabrication and analysis of deep submicron strained-Si N-MOSFET's. *IEEE Transactions on Electron Devices*, 47(7):1406 1415, 2000.

[154] K. Ismail, S. F. Nelson, J. O. Chu, and B. S. Meyerson. Electron transport properties of Si SiGe heterostructures: Measurements and device implications. *Applied Physics Letters*, 63(5):660 662, 1993.

[155] F. M. Buffer et al. Analysis of electron transport properties in unstrained and strained $Si_{1-x}Ge_x$ alloys. *Journal of Technology Computer Aided Design*, 1997.

[156] Private communication with Dr. T. Hackbarth, Daimler Chrysler, Germany.

[157] H. Rohdin. Reverse modeling of E D logic submicrometer MODFET's and prediction of maximum extrinsic MODFET current gain cutoff frequency. *IEEE Transaction On Electron Devices*, 37(4):920 933, 1990.

[158] P. J. Tasker and B. Hughes. Importance of source and drain resistance to the maximum f_T of millimeter wave MODFET's. *IEEE Electron Device Letters*, 10(7):291 293, 1989.

[159] J. Wood and C. G. Morton. An analysis of the effective position of the two-dimensional electron gas in the channel of MODFET epitaxial layer structures. *IEEE Transaction On Electron Devices*, 45(7):1622 1624, 1998.

[160] M. C. Foisy, P. J. Tasker, B. Hughes, and L. F. Eastman. The role of inefficient charge modulation in limiting the current-gain cutoff frequency of the MODFET. *IEEE Transactions on Electron Devices*, 35(7):871 878, 1998.

[161] G. Halkias and A. Vegiri. Device parameter optimization of strained si channel n-MODFET's using a one-dimensional charge control model. *IEEE Transactions on Electron Devices*, 45(12):2430 2436, 1998.

[162] T. Soorapanth and T. H. Lee. RF linearity of short-channel MOSFET. *First International Workshop on Design of Mixed-Mode Integrated Circuits and Applications*, pages 81 84, 1997.

[163] M.-T. Yang and Y.-J. Chan. Device linearity comparisons between doped-channel and modulation-doped designs in pseudomorphic $Al_{0.3}Ga_{0.7}As$ $In_{0.2}Ga_{0.8}As$ heterostructures. *IEEE Transactions on Electron Devices*, 43(8):1174 1180, 1996.

[164] Y. Taur, Y.-J. Mii, R. Logan, and H.-S. Wong. On "effective channel length" in 0.1-μm MOSFET's. *IEEE Electron Device Letters*, 16(4):136 138, 1995.

[165] T. Yaur and T. H. Ning. *Fundamentals of modern VLSI devices*. Cambridge University Press, 1998.

[166] K. L. Lee et al. Sub-30nm P+ abrupt junction formation in strained $Si/Si_{1-x}Ge_x$ MOS device. *Technical Digest of International Electron Devices Meeting*, pages 379 382, 2002.

[167] S. Thompson et al. Source/Drain externsion scaling for 0.1μm and below channel length MOSFETS. *Symposium on VLSI Technology Digest of Technical Papers*, pages 132 133, 1998.

[168] http://www-mtl.mit.edu/Well/.

[169] K. Kim, C.-T. Chuang, K. Rim, and R. V. Joshi. Performance assessment of scaled Strained-Si channel-on-insulator(SSOI) CMOS. *Proceedings of IEEE International SOI Conference*, pages 17 19, 2000.

[170] G. G. Shahidi. SOI technology for the GHz era. *IBM Journal of Research and Development*, 2002.

[171] 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 Transactions on Electron Devices*, 41(12):2357 2362, 1994.

[172] C. Jacoboni and L. Reggiani. The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials. *Reviews of Modern Physics*, 55(3):645 705, 1983.

[173] K. Kalna and A. Asenov. Nonequilibrium and ballistic transport, and backscattering in decanano HEMTs: a Monte Carlo simulation study. *Mathematics and Computers in Simulation*, 62:357 366, 2003.

[174] S. E. Laux, M. V. Fischetti, and D. J. Frank. Monte Carlo analysis of semiconductor devices: The DAMOCLES program. *IBM Journal of Research and Development*, 34(7):466 494, 1990.

[175] E. Sangiorgi and M. R. Pinto. A semi-empirical model of surface scattering for Monte Carlo simulation of silicon n-MOSFET's. *IEEE Transactions on Electron Devices*, 39(2):356 361, 1992.

[176] M. V. Fischetti and S. E. Laux. Monte Carlo study of electron transport in silicon inversion layers. *Physical Review B*, 48(7):2244 2274, 1993.

[177] D. K. Ferry and H. L. Grubin. Electrons in semiconductors: How big are they? *Proceedings of the 6th International Workshop on Computational Electronics*, pages 84 87, 1996.

[178] S. M. Goodnick andothers. Surface roughness at the Si(100)-SiO_2 interface. *Physical Review B*, 32:8171 8186, 1985.

[179] M. V. Fischetti, F. Gámiz, and W. Hänsch. On the enhanced electron mobility in strained-silicon inversion layers. *Journal of Applied Physics*, 92(12):7320 7324, 2002.

[180] Y. H. Xie et al. Semiconductor surface roughness: Dependence on sign and magnitude of bulk strain. *Physical Review Letters*, 73:3006–3011, 1994.

[181] S. Inaba et al. High performance 35nm gate length CMOS with NO oxynitride gate dielectric and Ni salicide. *IEEE Transactions on Electron Devices*, 49(12):2263–2270, 2002.