STRAIN EFFECTS IN LOW-DIMENSIONAL SILICON MOS AND ALGAN/GAN HEMT DEVICES

By

MEHMET ONUR BAYKAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2012

© 2012 Mehmet Onur Baykan

To my mother, my father, my brother and my baby nephew, Efe

ACKNOWLEDGMENTS

First of all, I would like to thank my advisor, Dr. Toshikazu Nishida, to whom I am indebted for his invaluable guidance and support throughout my entire graduate research. Also, I would like to thank my co-advisor, Dr. Scott Thompson, for his insightful contributions and inspirational perspective on research problems. I also would like to thank my Ph.D. committee members, Dr. Jing Guo, Dr. Ant Ural and Dr. Brent Gila for their suggestions and interests in my work.

In addition, I would like to thank Air Force Office of Scientific Research for their support of my graduate research through MURI project funding.

I also would like to thank my colleagues at Sematech, Dr. Chris Hobbs, Dr. Kerem Akarvardar, Dr. Chadwin Young, and Dr. Prashant Majhi for the opportunities during my internship in 2010-2011, which gave me a professional perspective of the state of the art semiconductor industry.

Also, I would like to thank all my lab mates: Amit, Andy, Ashish, David, Hyunwoo, Jingjing, Min, Nicole, Sagar, Sri, Tony, Ukjin, Uma, Xiaodong, Yongke and Younsung, for their help, support and friendship which makes everyday at work a joyful and a productive one.

I would like to give special thanks to The Scientific and Technological Research Council of Turkey (TUBITAK) for their generous support of my doctoral studies.

Finally, I would like to thank my mother, my father, my brother and my dear Helen for their everyday support and encouragement.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS 4

LIST OF TABLES 8

LIST OF FIGURES 9

ABSTRACT 13

CHAPTER

1 INTRODUCTION 16

1.1 Motivation 16
1.2 Objective and Organization 19

2 BACKGROUND ON CLASSIFICATION OF DEVICES IN TERMS OF CARRIER DEGREE OF FREEDOM 21

2.1 Motivation 21
2.2 Classification of Devices in Terms of Carrier Degree of Freedom 22
2.2.1 Unconfined/3DOF Bulk Devices 22
2.2.2 Devices with Predominant Electrostatic Confinement (EC) 25
2.2.2.1 1D EC/2DOF surface-inversion MOSFET 25
2.2.2.2 1D polarization-induced EC/2DOF 2DEG AlGaN/GaN HEMT devices 27
2.2.2.3 1D EC/2DOF volume-inversion MuGFET devices 28
2.2.2.4 2D EC/quasi-1DOF volume-inversion devices 29
2.2.3 Devices with Predominant Spatial Confinement 32
2.2.3.1 1D SC/2DOF ultra-thin-body (UTB) devices 32
2.2.3.2 2D SC/1DOF nanowires (NW) 34
2.3 Summary 38

3 BACKGROUND ON TRAP CHARACTERIZATION METHODS 39

3.1 Motivation 39
3.2 Overview of Trap Characterization Methods 40
3.2.1 Pulsed I-V 40
3.2.2 Conductance Dispersion 41
3.2.3 Subthreshold Swing 43
3.2.4 Current Transients 43
3.2.5 Deep Level Transient Spectroscopy (DLTS) 44
3.2.6 Photoionization Spectroscopy and Luminescence 45
3.3 Summary 46

4 INSIGHTS ON THE (100) AND (110) DOUBLE-GATE FINFET ELECTRON TRANSPORT AND THE STRAIN RESPONSE OF P-TYPE TRI-GATE FINFETS 48

4.1 Physical Insights on Comparable Electron Transport in (100) and (110) Double-gate Fin Field-effect Transistors . . . 48
4.1.1 Motivation . . . 48
4.1.2 Experimental . . . 49
4.1.3 Results and Discussion . . . 51
4.1.4 Conclusion . . . 54
4.2 Strain Response of p-type Tri-Gate FinFETs . . . 54
4.3 Summary . . . 58

5 SIZE AND ORIENTATION DEPENDENT STRAIN EFFECTS ON BALLISTIC SI P-TYPE NANOWIRE FIELD EFFECT TRANSISTORS . . . 60

5.1 Motivation . . . 60
5.2 Method . . . 61
5.3 Results and Discussion . . . 66
5.3.1 Unstrained Devices . . . 68
5.3.2 Strained Devices . . . 71
5.4 Summary . . . 75

6 PHYSICAL INSIGHTS ON STRAIN ENHANCED SURFACE ROUGHNESS LIMITED ELECTRON MOBILITY IN SILICON INVERSION LAYERS . . . 77

6.1 Motivation . . . 77
6.2 Experimental . . . 79
6.3 Simulation . . . 81
6.3.1 Temperature Dependence . . . 81
6.3.2 Strain Modification . . . 82
6.3.3 Mobility Model . . . 83
6.3.3.1 Intravalley phonon scattering . . . 84
6.3.3.2 Intervalley phonon scattering . . . 85
6.3.3.3 Surface roughness scattering . . . 86
6.4 Results and Discussion . . . 88
6.5 Summary . . . 97

7 UNIFIED QUASI-BALLISTIC COMPACT MODEL FOR STRAINED NANOSCALE MOSFETS USING AN UPDATED ONE-FLUX THEORY . . . 99

7.1 Motivation . . . 99
7.2 Fundamentals . . . 101
7.2.1 Natori's Original Ballistic Transport Model . . . 101
7.2.2 Current Control in Natori's Model . . . 103
7.2.3 One-flux Theory: Basics . . . 104
7.2.4 Revisiting the Lundstrom Model . . . 108
7.2.5 Current Control in Lundstrom Model . . . 112

7.2.6 Issues in the Lundstrom Model . . . 114
7.2.7 Revisiting the Gildenblat Model . . . 117
7.2.8 Revisiting the Original Shockley Work . . . 120
7.3 UF Compact Model: Theory . . . 121
7.3.1 Including High Field Effects . . . 122
7.3.2 Including Confinement Effects . . . 125
7.3.3 Revisiting the Natori's High Field Transport Model . . . 126
7.3.4 Understanding Feedback in Natori's High Field Model . . . 130
7.3.5 Physical Significance of Effective Velocity Term . . . 136
7.3.6 Expanding Natori's High Field Model for MOSFETs . . . 139
7.4 Qualitative Understanding of Strain Effects . . . 140
7.5 Summary . . . 142

8 OPTICAL CHARACTERIZATION OF ALGAN/GAN HEMT TRAP GENERATION 144

8.1 Motivation . . . 144
8.2 Overview of HEMT Device . . . 146
8.3 Method . . . 148
8.4 Experimental Setup . . . 150
8.5 Results and Discussion . . . 157
8.6 Summary . . . 165

9 SUMMARY AND RECOMMENDATION FOR FUTURE WORK . . . 166

9.1 Research Summary . . . 166
9.2 Recommendation for Future Work . . . 168

REFERENCES . . . 170

BIOGRAPHICAL SKETCH . . . 192

LIST OF TABLES

Table page

3-1 Summary of different trap characterization methods in terms of capabilities of density, time constant and activation energy extraction, and challenges. 47

5-1 The actual atomistic height and width of the nanowires used in this study. . . . 64

6-1 The values of phonon deformation potentials, phonon energies, and surface roughness parameters used for mobility calculations are shown. 90

LIST OF FIGURES

Figure page

1-1 The evolution of the cost of lithography tools with time. 17

1-2 The illustration of the spontaneous and piezoelectric polarization in AlGaN/GaN heterostructure is shown. 19

2-1 Classification of semiconductor devices based on the type and dimensionality of carrier confinement. 22

2-2 Various 2DOF silicon devices are shown. 23

2-3 Electrically confined quasi-1DOF volume-inversion and spatially confined 1DOF nanowire silicon MOSFET channels are shown. 24

2-4 Conduction band minima for non-spatially confined unstrained silicon devices are shown. 25

2-5 Spatially confined 2DOF UTB silicon and 1DOF silicon nanowire devices are shown. 31

2-6 The direct bandgap transition from 3DOF bulk silicon to to (001) 2DOF UTB Si (left) and $\langle 100 \rangle$ channel 1DOF SiNW (right) are illustrated. 33

2-7 The valence bands for unstrained $2.1nmX2.1nm$ silicon nanowires are shown. . 35

3-1 The pulsed I-V method for high-κ silicon MOSFET is shown. 41

3-2 Equivalent circuit models are shown for the conductance dispersion method. . 42

3-3 Photoemission is shown for pre-stress and post-stress conditions under forward and reverse bias. 46

4-1 Effective electron mobility is extracted from (100) and (110) sidewall FinFETs using split-CV method, and shown as a function of N_{inv}. 50

4-2 Flowchart of modified Schrödinger-Poisson self-consistent simulation procedure to take non-parabolic confinement mass into account. 51

4-3 Energy-wavevector dispersion relation of the transverse electron mass and the extracted confinement mass/quantization energy pairs are shown. 52

4-4 Evolution of (110) CB valley degeneracies is shown with and without non-parabolic confinement mass approach. 52

4-5 Average electron conductivity effective mass vs. N_{inv} is shown for both (100) and (110), using both parabolic and non-parabolic confinement mass. 53

4-6 Tri-gate p-type FinFET sample cleaved from a 300mm wafer is under ∼1GPa uniaxial tension. 55

4-7 4-point uniaxial wafer bending is illustrated. 56

4-8 The external parasitic resistance R_{SD} extracted from n (blue) and p-type (red) tri-gate FinFETs using shift and ratio method. 56

4-9 Experimentally extracted piezoresistance coefficients from p-type tri-gate FinFETs are shown. 57

5-1 Straight ⟨110⟩ and 45° rotated ⟨100⟩ nanowire devices are shown on conventional (001) silicon-on-insulator wafer. 62

5-2 For a constant number of mobile charges in a unit volume flowing along L, the equivalent 3D, 2D and 1D charge densities are illustrated. 67

5-3 The valence band structure is shown for unstrained narrowest ($3\mathrm{nm\ x\ 3nm}$) and widest ($7\mathrm{nm\ x\ 7nm}$) nanowires in both ⟨110⟩ and ⟨100⟩ directions. 68

5-4 The size and orientation dependent unstrained ballistic average injection velocity and hole drain current are shown. 70

5-5 The band structure hole distribution is shown for 3nm X 3nm ⟨100⟩ nanowire under 0.8% tension, unstrained and 0.8% compression at $|V_{ov}| = 0.6V$. 72

5-6 The band structure current distribution is shown for 3nm X 3nm ⟨100⟩ nanowire under 0.8% tension, unstrained and 0.8% compression at $|V_{ov}| = 0.6V$. 73

5-7 The band structure hole distribution is shown for 3nm X 3nm ⟨110⟩ nanowire under 0.8% tension, unstrained and 0.8% compression at $|V_{ov}| = 0.6V$. 74

5-8 The band structure current distribution is shown for 3nm X 3nm ⟨110⟩ nanowire under 0.8% tension, unstrained and 0.8% compression at $|V_{ov}| = 0.6V$. 74

5-9 Size dependent gauge factor is shown for both straight and rotated p-type nanowires at an overdrive voltage of 0.6V. 76

6-1 The cryogenic temperature wafer bending setup and bending apparatus are shown. 79

6-2 The strain enhancement of nMOSFET linear drain current is shown as a function of sample temperature for two different stress conditions. 80

6-3 AFM system coupled to mechanical bending jig and the measured roughness amplitude are shown. 80

6-4 Double-gate and single-gate operation of the FinFET simulator is demonstrated. 81

6-5 The effective mobility and its constituents for a single-gate bulk MOSFET are plotted vs. effective electric field at 300K and no stress. 89

6-6 Electron mobility vs. inversion electron density is shown for a bulk single-gate (100) MOSFET and a 20nm thick double-gate (100) MOSFET at 300K. 91

6-7 Electron mobility vs. effective electric field is shown for a bulk single-gate (100) MOSFET and a 20nm thick double-gate (100) MOSFET at 300K. 92

6-8 Simulated single-gate MOSFET low-field electron mobility vs. gate voltage is shown for a temperature range of 400K to 100K. 93

6-9 For the single gate MOSFET, simulated inversion carrier density vs. gate voltage is shown for temperatures ranging from 400K to 100K with decrements of 50K. 94

6-10 Simulated single-gate MOSFET linear drain current vs. gate voltage is shown for a temperature range of 400K to 100K. 94

6-11 Simulated single-gate nMOSFET normalized channel resistance change vs. uniaxial stress and the expected change from the π-coefficient are shown. . . . 95

6-12 The *ad hoc* simulation results are shown for the 75MPa condition. 96

6-13 Normalized *ad hoc* fitting error is shown for 25MPa, 50MPa and 75MPa conditions for a given gate voltage. 97

7-1 Schematic of conduction or first subband profile along the length of a MOSFET channel for device biased in saturation conditions. 102

7-2 Incoming and outgoing carrier fluxes are shown for a slab of thickness dx. . . . 106

7-3 Definition of the KT-layer in the Lundstrom Model, transmitted and backscattered carriers near the critical region for transport are shown. 109

7-4 Natori's high field transport model showing the conceptual "Initial Elastic Zone" and "Optical Phonon Emission Zone" regions in a device carrying current. . . . 128

7-5 Our derivation of the overall transmission coefficient using an infinite series summation approach. 132

7-6 Equivalent negative feedback system representations of the Natori Model are shown. 132

7-7 Evolution of MOSFET conduction band and channel transmission characteristics is shown for different transistor operation conditions. 136

8-1 The sample information is provided for the AlGaN/GaN HEMT structures used in this study. 146

8-2 Experimentally measured capacitance-voltage and extracted 2DEG density are shown for CV-dots on both Si and SiC substrates. 148

8-3 The carrier dynamics of the photoionization spectroscopy method are illustrated.149

8-4 The experimental setup used in this study is shown. Both electrical and mechanical degradation experiments are integrated into the PS measurements. 151

8-5 The linear drain current response from a 30μm long TLM device is shown for both with and without focusing lenses (i.e. high and low intensity) 650nm illumination.152

8-6 The transmission characteristics of the UV-rejection higher order filter is shown. 153

8-7 Cross calibration curves between Spectra-650 colorimeter and Oriel 70124 pyroelectric detector. 155

8-8 The normalized photon flux with respect to 360nm illumination vs. incident photon wavelength is shown. 156

8-9 The density normalization coefficient is shown for different photon wavelengths and energies. 157

8-10 The current-voltage curves are shown for 30 μm TLM structures on Si and SiC substrates. 158

8-11 A typical photoionization spectroscopy data from a 5μm TLM site on Si substrate is shown. 159

8-12 Extracted photon flux-normalized relative areal trap density is shown for a 30μm long TLM on SiC and a 5μm long TLM on Si substrate. 160

8-13 The repeatability of the PS experiment is demonstrated on a virgin 100nm off-center gate HEMT on SiC substrate. 161

8-14 Data collected from two consecutive runs of OFF-state electrical gate stress experiments is shown for a 100nm off-center gate HEMT on SiC substrate. . . 162

8-15 Gate current density is shown for pre-stress and post-OFF-state electrical stress conditions. 163

8-16 Width normalized drain current is shown for pre-stress and post-OFF-state electrical stress conditions. 164

8-17 Extracted photon flux-normalized relative electron and hole areal trap densities are shown for before and after the OFF-state stress. 164

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

STRAIN EFFECTS IN LOW-DIMENSIONAL SILICON MOS AND ALGAN/GAN HEMT DEVICES

By

Mehmet Onur Baykan

August 2012

Chair: Toshikazu Nishida
Cochair: Scott E. Thompson
Major: Electrical and Computer Engineering

Strained silicon technology is a well established method to enhance sub-100nm MOSFET performance. With the scalability of process-induced strain, strained silicon channels have been used in every advanced CMOS technology since the 90nm node. At the 22nm node, due to the detrimental short channel effects, non-planar silicon CMOS has emerged as a viable solution to sustain transistor scaling without compromising the device performance. Therefore, it is necessary to conduct a physics based investigation of the effects of mechanical strain in silicon MOS device performance enhancement, as the transverse and longitudinal device dimensions scale down for future technology nodes.

While silicon is widely used as the material basis for logic transistors, AlGaN/GaN HEMTs promise a superior device platform over silicon based power MOSFETs for high-frequency and high-power applications. In contrast to the mature Si crystal growth technology, the abundance of defects in the GaN material system creates obstacles for the realization of a reliable AlGaN/GaN HEMT device technology. Due to the high levels of internal mechanical strain present in AlGaN/GaN HEMTs, it is of utmost importance to understand the impact of mechanical stress on AlGaN/GaN trap generation.

First, we have investigated the underlying physics of the comparable electron mobility observed in (100) and (110) sidewall silicon double-gate FinFETs, which is

different from the observed planar (100) and (110) electron mobility. By conducting a systematic experimental study, it is shown that the undoped body, metal gate induced stress, and volume-inversion effects do not explain the comparable electron mobility. Using a self-consistent double-gate FinFET simulator, we have showed that for (110) FinFETs, an increased population of electrons is obtained for the Δ_2 valley due to the heavy nonparabolic confinement mass, leading to a comparable average electron transport effective mass for both orientations.

The width dependent strain response of tri-gate p-type FinFETs are experimentally extracted using a 4-point bending jig. It is found that the low-field piezoresistance coefficient of p-type FinFETs can be modeled by using a weighted conductance average of the top and sidewall bulk piezoresistance coefficients.

Next, the strain enhancement of p-type ballistic silicon nanowire MOSFETs is studied using $sp^3d^5s^*$ basis nearest-neighbor tight-binding simulations coupled with a semiclassical top-of-the-barrier transport model. Size and orientation dependent strain enhancement of ballistic hole transport is explained by the strain-induced modification of the 1D nanowire valence band density-of-states. Further insights are provided for future p-type high-performance silicon nanowire logic devices.

A physics based investigation is conducted to understand the strain effects on surface roughness limited electron mobility in silicon inversion layers. Based on the evidence from electrical and material characterization, a strain-induced surface morphology change is hypothesized. To model the observed electrical characteristics, we have employed a self-consistent MOSFET mobility simulator coupled with an *ad hoc* strain-induced roughness modification. The strain induced surface morphology change is found to be consistent among electrical and materials characterization, as well as transport simulations.

In order to bridge the gap between the drift-diffusion based models for long-channel devices and the quasi-ballistic models for nanoscale channels, a unified carrier transport

model is developed using an updated one-flux theory. Including the high-field and carrier confinement effects, a surface-potential based analytical transmission expression is obtained for the entire MOSFET operation range. With the new channel transmission equation and average carrier drift velocity, a new expression for channel ballisticity is defined. Impact of mechanical strain on carrier transport for both nMOSFETs and pMOSFETs in both linear and saturation regimes is explained using the new channel transmission definitions.

To understand the impact of mechanical strain on AlGaN/GaN HEMT trap generation, we have devised an experimental method to obtain the photon flux-normalized relative areal trap density distribution using photoionization spectroscopy technique. The details of the trap extraction method and the experimental setup are given. Using this setup, the trap characteristics are extracted for both ungated transmission line module (TLM) and gated HEMT devices from both Si and SiC substrates. The changes in the device trap characteristics are emphasized before and after electrical stressing. It is found through the step-voltage stressing of the AlGaN/GaN HEMT gate stack that the device degradation is due to the near bandgap trap generation, which are shown to be related to the structural defects in GaN.

CHAPTER 1
INTRODUCTION

1.1 Motivation

Silicon based metal-oxide-semiconductor field-effect transistors (MOSFET) have been the primary choice of mainstream integrated circuit (IC) technology for over 40 years. The continuous geometric scaling of planar silicon MOSFETs enabled ICs with superior performance, density, speed, cost, functionality and power figures compared to their predecessors. However, as the channel length of the planar MOSFETs shrink to deep submicron regime, the limitations due to the increased lithography costs have slowed down the geometric scaling (Figure 1-1). In the meantime, strain engineering has emerged as a cheaper solution to silicon MOSFET enhancement compared to the expensive geometric scaling at the 90nm technology node [1]. Process-induced strain has increased MOSFET carrier mobility through the reduction of conductivity effective mass and scattering rates [2]. Due to the scalability of the process-induced strain, strained silicon channels have been used in all technology nodes since its demonstration in 2002. With the further scaling of the strained planar MOSFETs, deleterious short-channel effects has led to increased off-state leakage and power dissipation due to the decreased gate control on the channel resistivity. To mitigate these short channel effects, higher doping densities are used to increase the gate control by higher vertical electric field in the channel, while sacrificing the carrier mobility. One solution has emerged to the short-channel effects and lowered carrier mobility in highly doped channels with the introduction of the multi-gate non-planar silicon CMOS design with thin undoped channels [3–6]. In fact, Intel has recently announced their tri-gate non-planar strained silicon CMOS baseline for the upcoming 22nm node [7].

With the continuous channel length scaling for increased transistor density and performance, the dominant mechanism for carrier transport has evolved from drift-diffusion for long channel devices to quasi-ballistic for ultra short channel devices

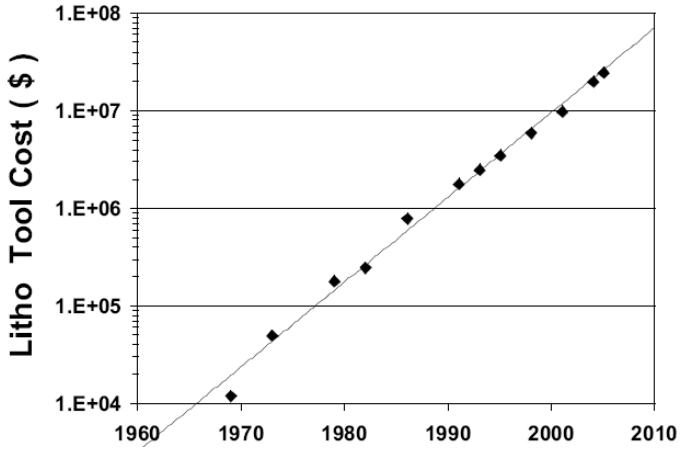


Figure 1-1. The evolution of the cost of lithography tools with time. Adapted from [8].

[9, 10], for which strain is expected to continue being a viable performance booster [11]. For the devices approaching the scaling limit beyond the 22nm node, carriers are expected to traverse the transistor channel at near ballistic transport conditions [12]. With the introduction of non-planar multi-gate silicon MOSFETs, a transverse dimension scaling is also required to control the short channel effects while decreasing the channel length at smaller nodes. Hence, ultra narrow strained gate-all-around silicon nanowire MOSFETs are expected to emerge to overcome the stronger short-channel effects in further scaled devices beyond the 22nm technology node. Consequently, there is a need for understanding the physics of strain enhancement of low-dimensional silicon MOS devices as the transverse and longitudinal dimensions scale down for future logic devices.

Although the silicon MOSFET technology is the mainstream for the state-of-the-art low-power and high-performance logic devices, silicon based power MOSFETs are outperformed by AlGaN/GaN high electron mobility transistors (HEMTs), especially for high-power and high-frequency applications [13]. The first AlGaN/GaN HEMT was demonstrated by Khan *et al.* in 1994 [14]. AlGaN/GaN HEMTs benefit from material

advantages over silicon power MOSFETs such as high breakdown voltage, due to its 3.4eV wide band gap, and lower capacitive loading at high frequencies due to smaller dielectric constant with a comparable heat dissipation [15, 16]. With these material advantages, the AlGaN/GaN HEMTs are shown to be capable of impressive high frequency and high power characteristics [17].

While mechanical stress is incorporated in silicon MOS device fabrication to improve performance with respect to its predecessors, a large built-in epitaxial stress in the AlGaN layer is required to realize the wide band gap GaN based HEMTs [18]. By epitaxially growing a thin AlGaN layer on a fully relaxed GaN substrate, the in-plane biaxial tension ($\sim 3GPa$) in the AlGaN barrier creates a fixed piezoelectric induced polarization charge added to the difference between AlGaN and GaN spontaneous polarization (Figure 1-2). The fixed positive polarization induced charge enables the accumulation of a two-dimensional electron gas (2DEG) via the donor states on the AlGaN layer [19]. In addition to the built-in epitaxial stress, the stress generated by the inverse piezoelectric effect in the AlGaN layer at high gate electric fields is shown to be on the order of 500MPa under high power operating conditions [20]. Clearly, mechanical strain is an essential component of the AlGaN/GaN HEMT operation.

Despite the impressive high frequency and high power performance, AlGaN/GaN HEMTs are far from reaching their full potential due to detrimental reliability issues [21]. Point and structural defects in the bulk GaN region, in the AlGaN barrier, at the AlGaN/GaN interface, at the Ni/AlGaN gate metal interface, and at the ungated region of the AlGaN surface create trapping centers [21–23]. These traps lead to degradation in the device performance through various mechanisms such as current collapse, drain lag and gate lag [21–25]. Moreover, the stress due to the inverse piezoelectric effect is hypothesized to lead to structural defects in the AlGaN barrier [20, 26]. All of these defects results in a challenge to achieve the realization of reliable AlGaN/GaN HEMTs for long-term applications. Since mechanical stress is inherently incorporated

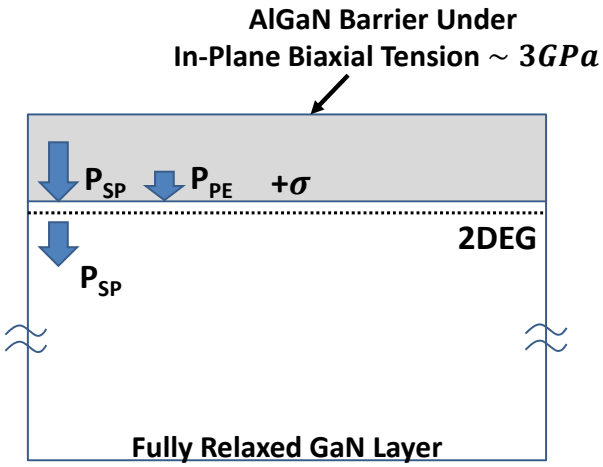


Figure 1-2. The illustration of the spontaneous and piezoelectric polarization in AlGaN/GaN heterostructure is shown. The net fixed positive charge (σ) on the AlGaN side of the interface creates a two-dimensional electron gas on the GaN side of the boundary.

in AlGaN/GaN HEMT operation, it is required to understand the impact of mechanical strain on AlGaN/GaN HEMT degradation and reliability.

1.2 Objective and Organization

The aim of this work is to understand the strain effects on the performance of the contemporary and future low dimensional silicon MOS devices, as well as the impact of the mechanical strain on AlGaN/GaN HEMT device reliability. To achieve this goal, a systematic approach of physics based models, appropriate simulation methods and wafer bending experiments were conducted. The contributions of this work include predicting the strain enhancement of future strained silicon MOSFETs through an improved understanding of contemporary devices and identifying the effects of mechanical strain on AlGaN/GaN HEMT device trap generation, degradation and reliability. These results may be used to guide paths for choosing device designs for future Si logic technology nodes and improved AlGaN/GaN HEMT designs for achieving reliable long-term operation.

In the next two chapters (2 and 3), we will first discuss the classification of the silicon MOS devices in terms of carrier degree of freedom. The device geometry dependent effects of electrostatic and spatial confinement on the silicon band structure are provided. Next, an overview of the widely used trap characterization methods will be presented from the perspective of their application to AlGaN/GaN devices, while emphasizing the merit of each technique.

Chapter 4 presents the physical insights on the experimentally observed comparable electron mobility in (100) and (110) sidewall silicon double-gate FinFETs. Supported with a systematical set of experiments, a possible physical mechanism is discussed as the result of a modified self-consistent Schrödinger-Poisson double-gate FinFET simulation. Then, the width dependence of the strain response of p-type tri-gate FinFETs is investigated with 4-point bending experiments and weighted average of bulk piezoresistive coeffficients.

In Chapter 5, the strain enhancement of ballistic silicon nanowire hole transport will be presented. The modification of the strain-induced ballistic current improvement as a function of nanowire channel direction and size is studied using an $sp^3d^5s^*$ basis nearest-neighbor tight-binding simulations.

Physical insights on the strain enhancement of surface roughness limited mobility in silicon inversion layers are presented in Chapter 6. The fundamental physics are investigated using an *ad hoc* simulation method supported by a combination of electrical and materials characterization.

Chapter 7 demonstrates the development and the details of a unified quasi-ballistic compact model for strained nanoscale MOSFETs using one-flux theory.

The optical characterization of strain effects on AlGaN/GaN HEMT trap generation will be presented in Chapter 8. Details on the samples, experimental setup, experiments and results will be discussed.

Chapter 9 provides the overall summary and future work.

CHAPTER 2
BACKGROUND ON CLASSIFICATION OF DEVICES IN TERMS OF CARRIER DEGREE OF FREEDOM

2.1 Motivation

Classification of devices as 3DOF, 2DOF, and 1DOF focuses on the effects of the predominant type of transverse quantum confinement (unconfined, electrostatic or spatial confinement) and device geometry (planar or non-planar) on the degrees of freedom of the carrier momentum. In addition to the transverse dimensions of the devices (confinement), effects of the longitudinal dimension of the device (long/short channel) will be considered from the perspective of the dominant carrier transport characteristics. The classification of semiconductor devices based on carrier confinement is shown in Figure 2-1. First, devices are divided into two high-level groups, unconfined and confined devices, based on the degrees of freedom of the carrier momentum. Unconfined devices refer to 3DOF bulk devices such as piezoresistors while confined devices include both devices with 1-dimensional confinement/2DOF carrier momentum and 2-dimensional confinement/1DOF carrier momentum. Next, confined 2DOF (Figure 2-2) and 1DOF (Figure 2-3) devices are divided into electrically confined (EC) and spatially confined (SC) devices. The electrically confined devices may be further subdivided in terms of semi-infinite substrate 2DOF planar MOSFET, 2DOF AlGaN/GaN HEMT, undoped 2DOF volume-inversion double-gate FET, and quasi-1DOF volume-inversion gate-all-around FET. Likewise, the spatially confined devices include 2DOF ultra-thin-body (UTB) silicon and 1DOF nanowire silicon (SiNW).

Parts are reprinted with permission from Mehmet O. Baykan, Scott E. Thompson, and Toshikazu Nishida, “Strain effects on three-dimensional, two-dimensional, and one-dimensional silicon logic devices: Predicting the future of strained silicon”, J. Appl. Phys. 108, 093716 (2010), DOI:10.1063/1.3488635. Copyright 2010, American Institute of Physics.

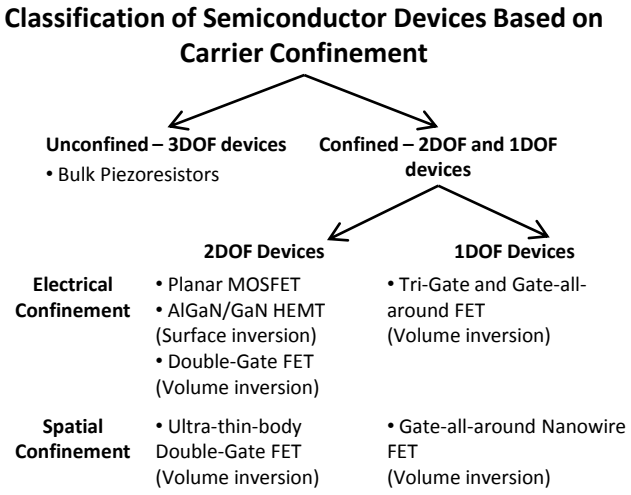


Figure 2-1. Classification of semiconductor devices based on the type and dimensionality of carrier confinement. Carriers in the semiconductor lattice obey the "Universal Rule of Dimensionality" for which $n + m = 3$, where n is the degree of freedom (DOF) of the carrier momentum and m is the dimensionality of confinement. In this sense, 3DOF carriers have no confinement (0C), 2DOF carriers have 1D confinement (1C) and 1DOF carriers have 2D confinement (2C). ©2010, American Institute of Physics.

In the following sub-sections, the electrical characteristics of unstrained 3DOF to 1DOF devices are reviewed in the framework of the classification scheme. We will discuss the effects of carrier confinement on the unconfined 3DOF bulk semiconductor band structure for confined devices by means of the effective mass of carriers, degeneracy of bands, and energy splitting between the lowest conduction and valence bands. The devices are grouped as: (A) Unconfined 3DOF Bulk Devices, (B) Electrically Confined 2DOF/quasi-1DOF FETs and (C) Spatially Confined 2DOF UTB and 1DOF Nanowire silicon devices.

2.2 Classification of Devices in Terms of Carrier Degree of Freedom

2.2.1 Unconfined/3DOF Bulk Devices

An example of a predominantly unconfined bulk device is a resistor where the majority carriers possess three degrees of freedom and are thus unconfined except at the boundaries. Hence, most of the mobile carriers do not interact with a surface, and carrier transport can be assumed to be free of any surface effects. A notable

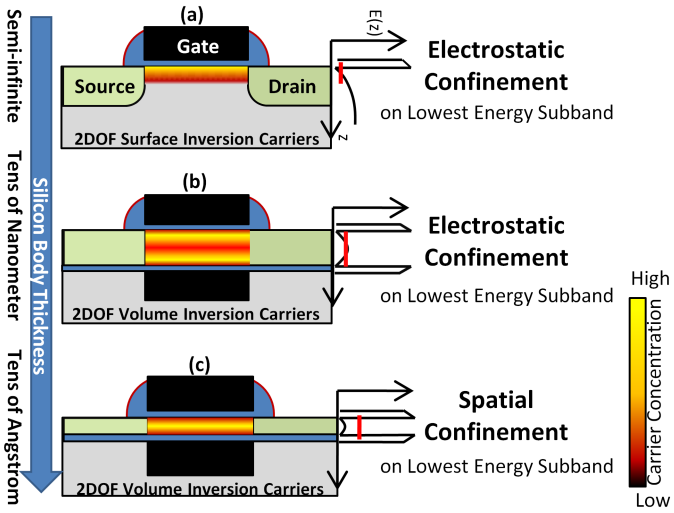


Figure 2-2. Various 2DOF silicon devices are shown. The carrier distribution inside the device channels is shown along with the conduction band edge profile for gate-oxide / silicon-channel heterostructure (black) and the lowest energy subband (red). (a) Electrically confined 2DOF planar MOSFET has surface-inversion carriers confined in the electrical well due to gate bias induced band bending. (b) For electrically confined 2DOF double-gate volume-inversion FET, carriers populate the bulk parts of the silicon, as the wavefunctions from the front and back gates interact. The lowest energy subband still lies in the electrical well defined by the gate bias induced band bending. (c) Spatially confined 2DOF UTB double-gate FET has the lowest energy subband confined in the silicon thickness defined spatial quantum well, and the peak of spatial carrier distribution occurs in the middle of the silicon slab. ©2010, American Institute of Physics.

strain-sensitive 3DOF device is a piezoresistor, a stress-dependent resistor widely used as a strain transducer in MEMS (Micro-Electro-Mechanical Systems). These devices transduce strain in the mechanical domain by exhibiting a strain-dependent resistance which is converted into a voltage signal by biasing the piezoresistor in a bridge circuit. Piezoresistance coefficients in 3DOF bulk silicon and germanium were first reported by Smith in 1954 [27].

As a baseline for comparison with confined devices, the band structure of unconfined 3DOF bulk silicon is first reviewed. In an unconfined bulk semiconductor, the carrier momentum is continuous in all three orthogonal directions in k-space. Neglecting

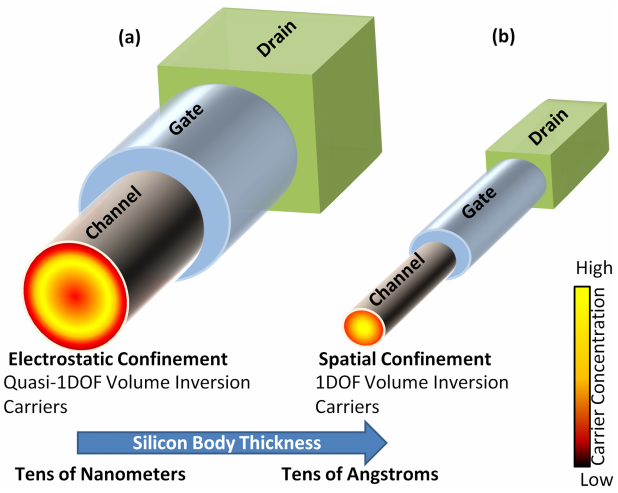


Figure 2-3. (Color Online) Electrically confined quasi-1DOF volume-inversion and spatially confined 1DOF nanowire silicon MOSFET channels are shown. Similar to their 2DOF counterparts, stronger volume-inversion is seen in these devices, as the predominant type of carrier confinement changes from electrical to spatial. ©2010, American Institute of Physics.

the conduction band edge energy and k-value at the minima, the kinetic energy of an electron under the parabolic band approximation is given by,

$$E = \frac{\hbar^2 k_x^2}{2m_x^*} + \frac{\hbar^2 k_y^2}{2m_y^*} + \frac{\hbar^2 k_z^2}{2m_z^*}, \tag{2–1}$$

where m_x^*, m_y^*, and m_z^* are the effective masses in the three directions. The conduction band (CB) of unconfined 3DOF bulk silicon consists of six-fold degenerate (Δ_6) prolate spheroidal equienergy surfaces along six equivalent Γ-X directions (Figure 2-4), with minima at $\pm 0.83(2\pi/a)$. For each Δ_6 energy-wave vector (E-k) valley, the mass along the major axis is $m_l = 0.916 m_o$ (longitudinal mass) and the masses along the minor axes are $m_t = 0.191 m_o$ (transverse mass), where m_o is the electron rest mass. Six-fold degeneracy of the CB edge implies equal electron occupancy of the Δ_6 valleys.

The valence band can be approximated by the highly anisotropic heavy-hole (HH) band, the relatively isotropic light-hole (LH) band, and the split-off (SO) band, all having maximum at the Γ point. The HH band exhibits the heaviest mass along the $\langle 110 \rangle$ direction with $m^*_{HH,\langle 110 \rangle} = 0.61 m_o$, while the LH band is fairly isotropic with $m^*_{LH} = 0.15 m_o$

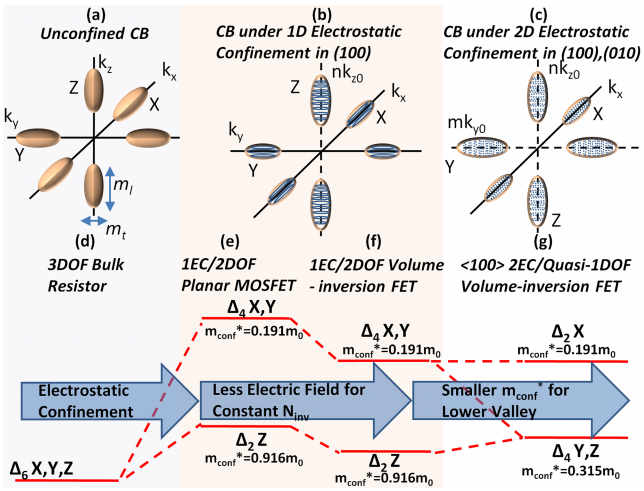


Figure 2-4. Conduction band (CB) minima for non-spatially confined unstrained silicon devices are shown. Each isoenergy ellipsoid has longitudinal mass $m_l = 0.92m_0$ and transverse mass $m_t = 0.19m_0$. (a,d) Unconfined 3DOF bulk silicon CB has 6-fold degenerate ellipsoid valleys in the 3D Brillouin Zone. (b) 1D confined (1C) band structure has quantized k_z (nk_{z_0}) and the corresponding 2D isoenergy contours denotes the CB edges for electrically confined 2DOF planar MOSFET (e) and 2DOF volume-inversion FET (f). (c) 2D confined (2C) CB has both k_y and k_z wavevectors quantized. The isoenergy ellipsoids are reduced to isoenergy dots due to the quantization in both y and z directions, and denote the CB edges for an electrically confined quasi-1DOF volume-inversion device (g). (d-g) The energy levels and shifts in CB edges due to electrical confinement is also shown for different devices with their contributor valleys, valley confinement masses and levels of degeneracy. Arrows explain the underlying physics for the changes in CB valley splitting during transition from one device to the other. ©2010, American Institute of Physics.

[2]. In unstrained 3DOF bulk silicon, the HH and LH bands are degenerate at the Γ point, while the SO band maximum is approximately 44meV lower than the HH and LH bands.

2.2.2 Devices with Predominant Electrostatic Confinement (EC)

2.2.2.1 1D EC/2DOF surface-inversion MOSFET

As shown in Figure 2-2-a, the surface potential induced transverse band bending of the conduction band edge (valence band edge) in n-type (p-type) inversion channel metal-oxide-semiconductor field-effect transistors (MOSFETs) fabricated on bulk silicon substrate confines electrons (holes) to a very narrow surface layer quantum well,

between the oxide potential barrier and the silicon band edge. Provided that the width of this quantum well is comparable to the de Broglie wavelength of the electron, the electron wavevector normal to the Si/SiO_2 interface is no longer continuous, but has discrete values. This quantization in the wavevector leads to formation of isoenergy band contours as shown in Figure 2-4-b. Although the mobile carriers are still allowed to freely move within the plane under the gate, motion of these carriers is restricted in the direction of confinement. In this case, the energy of a charge carrier is a function of the two continuous wavevectors and the discrete energy level defined by the quantized wavevector along the confinement direction. Assuming that the gate bias-induced surface electric field confinement occurs along the z-axis, the E-k dispersion relation can be defined by,

$$E_n = E_{z_n} + \frac{\hbar^2 k_x^2}{2m_x^*} + \frac{\hbar^2 k_y^2}{2m_y^*}, \tag{2–2}$$

where the energy of the nth inversion layer subband, E_{z_n}, under the depletion assumption is approximated by [28],

$$E_{z_n} = [\frac{3hq\varepsilon_s}{4\sqrt{2m_z}}(n + \frac{3}{4})]^{\frac{2}{3}}. \tag{2–3}$$

Here ε_s,m_z and E_{z_n} are the surface electric field, confinement mass, and discrete energy levels due to the quantized wavevector in the z direction respectively [28]. As seen in (2–3), the discrete energy levels are functions of the effective mass of the valley along the confinement direction and the surface electric field. This implies different subband energy levels for valleys of different confinement mass as shown in Figure 2-4-c.

Compared to unconfined 3DOF bulk silicon, inversion carriers in electrically confined 2DOF surface-inversion MOSFETs experience a modified band structure due to the gate-induced surface confinement electric field (ε_s), which shifts the energy levels of the valleys differently depending on their confinement effective mass. This

lifts the degeneracy and changes the occupation of the respective valleys. On (001) surface wafers, unlike the large difference between m_t and m_l of the conduction band, the out-of-plane masses of the HH and LH valence bands are very close to each other, $0.29m_o$ and $0.20m_o$ respectively [2]. Hence, at the inversion charge densities of contemporary 2DOF planar CMOS devices ($\sim 10^{13}cm^{-3}$), the energy splitting between the lowest subbands of the two ground-most valleys on (001) substrate are ~95meV for nMOS [29] but only ~20meV for pMOS [2]. As a result, in unstrained planar 2DOF pMOSFET devices, both the HH band (top-most) and the LH band (second) contribute to carrier transport, unlike the mostly isolated Δ_2 valley in n-type devices. In contrast, for (110) substrate, the confinement mass values for the HH and LH bands are $0.58m_o$ and $0.15m_o$ respectively [30], whereas, the confinement mass for Δ_2 is m_t and for Δ_4 is $m_{\langle 110\rangle} = 0.315m_o$ [31], possibly leading to a slightly greater band splitting of VB than CB.

2.2.2.2 1D polarization-induced EC/2DOF 2DEG AlGaN/GaN HEMT devices

Similar to 1D EC/2DOF surface-inversion silicon MOSFETs, the 1D polarization induced EC/2DOF AlGaN/GaN HEMT devices exhibit surface inversion layer characteristics. As briefly discussed in 1.1, the total interface fixed charge due to the spontaneous and piezoelectric polarization of the AlGaN/GaN heterostructure results in a quantum well created by the large conduction band bending at the AlGaN/GaN interface. As a result, a high density of two-dimensional electron gas accumulates in the high electron mobility GaN layer (Figure 1-2). In contrast to the enhancement mode operation of the 1D EC/2DOF silicon MOSFETs, the AlGaN/GaN HEMT is a depletion mode device with low resistivity (high inversion density) state at 0V gate bias. The 2DEG is depleted by applying a negative voltage on the gate below the threshold value.

Unlike the six-fold degenerate conduction band of silicon, the GaN conduction band is composed of a single isotropic minimum at the Γ point, with electron mass of $0.20m_0$ [32]. Therefore, the electrostatic confinement effect of valley degeneracy lifting is not seen for GaN based devices.

2.2.2.3 1D EC/2DOF volume-inversion MuGFET devices

The fundamental difference between conventional planar 2DOF surface-inversion MOSFET devices and volume-inversion 2DOF devices lies in the thickness of the silicon layer. Volume-inversion devices have a very thin silicon layer between two equipotential gates with a thickness ($\sim 20nm$) comparable to the cumulative bulk depletion width from each gate surface. When the silicon thickness is in this range, the electron wavefunctions confined by the two gates overlap spatially. As the magnitude squared of the wavefunction gives the carrier probability distribution in the device cross section, the carriers start to populate not only the near Si/SiO_2 interfacial region, but also the silicon region away from the gate. As shown in Figure 2-2-b, when the overlap of the wavefunctions is significant, the center volume of the thin silicon layer becomes inverted, resulting in a significant amount of inversion charge away from the gate surfaces. An increased number of inversion carriers in the center of the silicon slab reduces the band bending relative to the Si/SiO_2 interface, and hence reduces the overall confining effective electric field. The first 2DOF volume-inversion device was demonstrated by Balestra *et al.* [33]. Generally, the double-gate MOSFETs (DGFET) fall under this category. Another difference between the electrically confined 2DOF planar MOSFET and the 2DOF volume-inversion DGFET is in the doping of the silicon substrate. In contrast to the planar MOSFET, the channels of volume-inversion devices are mostly undoped and have less surface electric field than 2DOF planar MOSFETs. As (2–3) suggests, a reduced surface electric field in 2DOF volume-inversion devices leads to a smaller subband splitting compared to conventional 2DOF planar bulk substrate devices, as illustrated in Figure 2-4-f. Additionally, the wavefunctions that stem from the equipotential gates create overlapping energy bands. Although these overlapping bands split due to a mechanism similar to the Pauli's Exclusion Principle, the energy difference is generally smaller than $k_B T$ and treated as equally occupied [30]. Under this condition, a larger density-of-states for the ground subband that can

be occupied by a higher density of mobile inversion charges is obtained in 2DOF volume-inversion MuGFETs compared to 2DOF planar bulk substrate MOSFETs.

As the silicon layer thickness approaches the predominantly spatial confinement limit of $\sim 4nm$, further changes are observed in the band structure of 2DOF volume-inversion devices compared to conventional 2DOF planar surface-inversion devices. When the thickness of the silicon layer becomes comparable to the width of the wider, second lowest energy valley/band; the energy of the second lowest energy band increases due to spatial confinement, resulting in an increase in the splitting between the two ground-most valleys/bands. For the conduction band, the partial spatial confinement increases the energy level of the *primed* ladders (nomenclature is after Stern *et al.* [34]). Since the *unprimed* valleys/subbands have smaller width, electrical confinement remains as the main source of carrier confinement on the unprimed valleys, thus no energy increase is observed for the unprimed subbands under partial spatial confinement [35]. The result is increased valley splitting while still having an increased density-of-states in the ground subband due to wavefunction overlap. Under this condition, the ground subband, already occupied at a high level, increases its occupation further due to the onset of spatial confinement of the primed valleys/subbands. Likewise for the valence band of partially spatially confined silicon, holes mostly reside in the top-most (HH) band close to the Γ point due to increased band splitting and increased DOS of the ground subband.

2.2.2.4 2D EC/quasi-1DOF volume-inversion devices

Thick nanowire (widths between $\sim 4 - 20nm$), Tri-Gate FinFET and gate-all-around FinFET (GAA-FET) devices are considered as quasi-1DOF volume-inversion devices. The thickness range is $\sim 4 - 20nm$ since below approximately $4nm$, spatial confinement becomes dominant while above approximately $20nm$, wavefunction overlap is negligible. Extending the discussion on 2DOF volume-inversion devices to their quasi-1DOF counterparts, we observe that gate-all-around devices demonstrate similar characteristics

of volume-inversion, such as decreased band splitting and increased DOS of the ground valley/subband. However, the difference in device electrostatics is significant in 2DOF and quasi-1DOF volume-inversion devices. In a 1D EC/2DOF-surface device with a planar gate, the potential variation occurs primarily along one axis, resulting in 1D electrostatics. The discrete energy levels approximated by (2–3) are only dependent on the 1D electric field and mass along the electrostatic confinement direction. While planar 2DOF devices have gate(s) only in one plane and a 1D electric field, gate-all-around 1DOF-volume devices have surrounding gates and a 2D electric field. As seen in Figure 2-4-c, with the quantization in both k_y and k_z wavevectors, isoenergy band surfaces of 3DOF bulk silicon reduce to isoenergy band dots. The interaction of the Y and Z conduction band ellipsoids of the $\langle 100 \rangle$ channel quasi-1DOF volume-inversion device with the 2D gate electric field involves both the longitudinal and transverse silicon effective mass. Under the assumption that the surrounding gates are equipotential and create symmetric 2D electric field contours as illustrated in Figure 2-3-a, the energy shift of the Y and Z CB valleys (E_{Y_n} and E_{Z_n}) can be approximated by using (2–3). Under the isotropic confinement mass approximation, the effective mass for valley ν along the y and z directions is given by [36]

$$\frac{1}{m^*_{\nu,conf}} = \int_0^{2\pi} \frac{d\phi}{2\pi} \left(\frac{cos^2\phi}{m_{\nu,y}} + \frac{sin^2\phi}{m_{\nu,z}} \right), \tag{2–4}$$

where $m_{\nu,y}$ and $m_{\nu,z}$ are effective mass values for valley ν along y and z directions and ϕ is the positive angle from k_y-axis to k_z-axis [36].

Now, the fundamental difference between 2DOF and quasi-1DOF volume-inversion devices becomes apparent. While the electric field induced valley splitting in a (001) 2DOF device is implied by confinement masses of $m_t = 0.191 m_o$ (for X,Y valleys) and $m_l = 0.916 m_o$ (for Z valley), these mass values change to $m_t = 0.191 m_o$ (for X valley) and $0.315 m_o$ (for Y,Z valleys) for a $\langle 100 \rangle$ channel quasi-1DOF volume-inversion device. Similarly for $\langle 110 \rangle$ channel quasi-1DOF volume-inversion devices, we have average

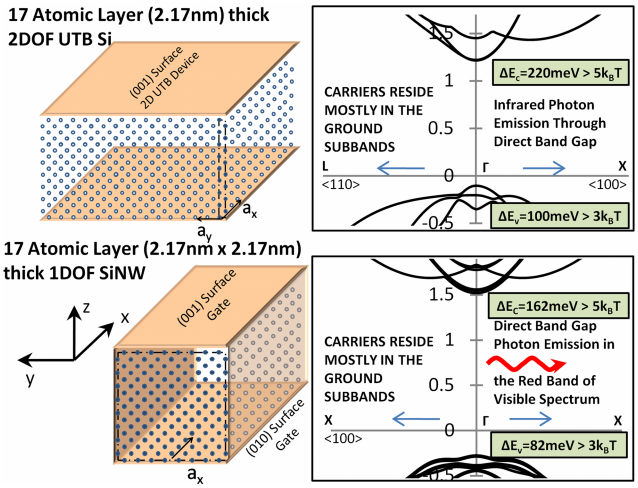


Figure 2-5. Spatially confined 2DOF UTB silicon and 1DOF silicon nanowire devices are shown in the figure. Silicon atoms in dark color and in frames are constructed unit cells for $sp^3d^5s^*$ tight-binding method. In accordance with the real lattice periodicity of unit cells for 2DOF UTB and 1DOF SiNW, real lattice vectors a_z in UTB and a_y,a_z in SiNW are not defined. The result is undefined k_z for 2DOF UTB and k_y,k_z for 1DOF SiNW reciprocal lattice. Bulk-Si valleys along the undefined reciprocal vectors are projected to origin of Brillouin Zone (Γ point), converting material from indirect to direct bandgap characteristics. Inset shows simulated band structure of devices and corresponding energy splitting levels and direct band gap energies [38]. High valley/band splitting in these devices ensures that carriers reside in the ground most subband, and carrier optical phonon scattering is suppressed. ©2010, American Institute of Physics.

confinement mass values of $0.237m_o$ (for X,Y valleys) and $0.315m_o$ (for Z valley). Accordingly when we consider (2–3), a smaller difference in confinement effective masses of CB valleys in a quasi-1DOF volume-inversion device is expected to result in a smaller confinement induced CB splitting than 2DOF surface and volume-inversion devices, as illustrated in Figure 2-4-g.

It should be pointed out that multi-gate devices having thicknesses greater than the cumulative depletion width (no spatial wavefunction/inversion-layer overlap), are essentially comprised of multiple independent channels [37]. Hence, these devices may be considered as multiple 2DOF planar surface-inversion MOSFETs operating in parallel with each other and do not exhibit the volume-inversion characteristics described above.

2.2.3 Devices with Predominant Spatial Confinement

When one or more dimensions of a semiconductor slab are scaled down to be comparable to the Bohr exciton radius of the material, spatial confinement of the carriers reduces the degrees of freedom of the carrier momentum and hence the reciprocal space, leading to quantization of the carrier momentum and splitting of the carrier energy. Illustrated in Figure 2-2-c and Figure 2-3-b, the volume-inversion effect is strongly present in these devices due to the very strong interaction of the multiple gates. For these spatially confined silicon devices, exceptionally different electrical and optical properties compared to non-spatially confined silicon are predicted and observed [39–44]. For spatially confined 2DOF ultra-thin-body (UTB) and 1DOF nanowire (NW) devices, the finite silicon crystal that constitutes the transistor channel is formed by a "countable" number of atomic layers, as shown in Figure 2-5. The atomic nature of these devices implies that first principles calculations are required to capture the unique electrical characteristics of these novel devices. The $sp^3d^5s^*$ tight-binding method (TB) and density functional theory (DFT) are the two most common methods employed to calculate the atomic band structure. These methods capture the band structure effects of these novel devices more accurately than the conventional effective mass approximation (EMA) [45]. The E-k dispersion relations for 2DOF UTB and 1DOF SiNW devices are calculated by TB method [38] and shown in Figure 2-5. Qualitatively, the crucial band structure properties that define the electrical and optical characteristics, such as the effective mass, valley/band energy splitting, and location of band extrema show remarkable differences compared to non-spatially confined silicon.

2.2.3.1 1D SC/2DOF ultra-thin-body (UTB) devices

Shown in Figure 2-5, 2DOF UTB Si devices are extremely confined in one direction and are assumed to possess an unperturbed lattice periodicity along the in-plane directions. The UTB silicon layer can be constructed from a single unit cell, extending from the bottom gate to the top gate and by repeating the unit cell along the unit real

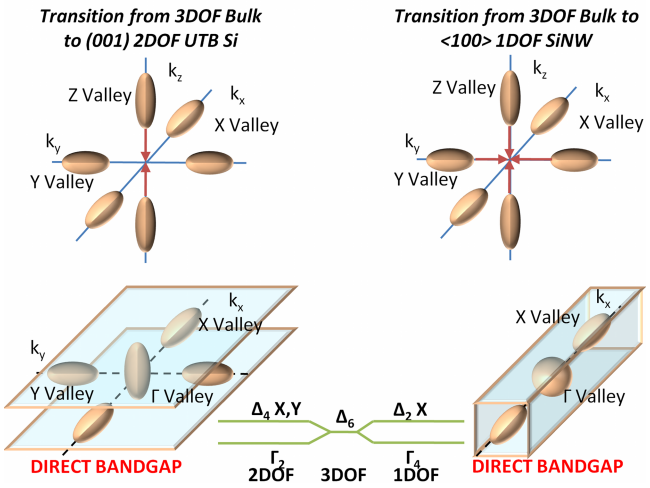


Figure 2-6. The direct bandgap transition from 3DOF bulk silicon to to (001) 2DOF UTB Si (left) and ⟨100⟩ channel 1DOF SiNW (right) are illustrated. The Z-valleys for UTB and the Y, Z-valleys for SiNW are shown as they are projected on to the origin of first Brillouin Zone by red arrows. CB edges at the Γ point are less upshifted by confinement effects than other valleys due to their higher confinement effective mass. These results of spatial confinement and confinement directions lead to direct band gap characteristics. ©2010, American Institute of Physics.

lattice vectors $\hat{a}_x$ and $\hat{a}_y$. This approach inherently leads to a 2D reciprocal space with wavevectors only along k_x and k_y directions. The wavevector, k_z, is undefined and its E-k dispersion is projected onto the k_x-k_y plane with $k_z = 0$, i.e. $E(k)_{UTB} \approx E(k_x, k_y, k_z)_{|k_z=0}$, within the 1st Brillouin Zone (BZ) [43]. This implies a special case for (001) surface devices where the bulk Z-valley at [0, 0, $\pm$0.83]($2\pi/a_0$) of the 3D BZ appears at the Γ point of the 2D BZ (Figure 2-6) as a Γ valley. For other surface orientations, none of the conduction band edges collapse onto the Γ point of the 2D BZ due to the presence of non-zero components of the in-plane wavevector in the coordinates of the bulk CB minima.

As illustrated in Figure 2-6, spatial confinement of (001) surface 2DOF UTB silicon device results in a direct band gap band structure, unlike the indirect band gap of non-spatially confined silicon. For (001) surface UTB Si, the off-Γ valleys originate from the X and Y bulk-valleys. The off-Γ valleys have smaller confinement mass and

larger transport mass. On the other hand, the aforementioned Γ valley conversely has a larger confinement mass and a smaller transport mass. Consequently, higher levels of quantum confinement (achieved by either decreasing the silicon thickness or increasing the gate voltage) further split the four-fold off-Γ valleys and two-fold Γ valleys. The result is a smaller upshift of the Γ valley compared to the off-Γ valley. When the Γ valley becomes the lowest edge of the (001) 2DOF UTB conduction band, an indirect to direct band gap transition occurs [43]. Similarly, the valence bands are downshifted in energy with increasing confinement in accordance with their confinement masses. Thus, an increase in forbidden energy band gap is observed with increased spatial confinement. The band splitting energies and direct bandgap for a 2.1nm thick (001) UTB Si device calculated using tight-binding method are shown in Figure 2-5 [38].

Finally, high levels of spatial confinement seen in 2DOF UTB devices result in deviation from the widely accepted m_l and m_t mass values of the CB electrons, as well as the VB hole effective masses. The underlying physics for this observation (change in curvature of band edges) will be discussed along with a similar discussion on 1DOF SiNWs within the next subsection.

2.2.3.2 2D SC/1DOF nanowires (NW)

While 2DOF UTB devices are spatially confined in one dimension, carriers in silicon nanowire (SiNW) devices possess a single DOF if the carriers are spatially confined in two dimensions. As shown in Figure 2-5, the unit cell for a 1DOF SiNW extends only along the $\hat{a}_x$ direction. Due to the spatial confinement in the transverse directions, for a $\langle 100 \rangle$ channel SiNW, only one wavevector in the k_x direction is defined. The $\langle 100 \rangle$ SiNW unit cell has twice the size of a bulk-Si unit cell along the direction of periodicity (x-axis) [46]. While the bulk Si BZ boundaries occur at $\pm 2\pi/a_0$, where a_0 is the bulk-Si lattice constant, the $\langle 100 \rangle$ SiNW 1st BZ occurs at $\pm \pi/a_0$ [46].

Similar to the 2D (001) UTB, the Y and Z bulk-valleys are projected to the Γ point, resulting in a direct band gap semiconductor for spatially confined $\langle 100 \rangle$ SiNW (Figure

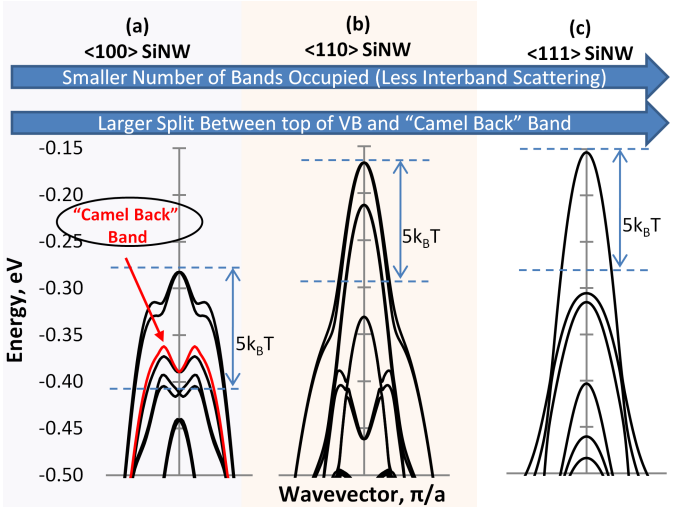


Figure 2-7. The valence bands for unstrained $2.1nmX2.1nm$ thick (a) $\langle 100 \rangle$, (b) $\langle 110 \rangle$ and (c) $\langle 111 \rangle$ square silicon nanowires are simulated through online resources [38]. At room temperature, smaller number of difference valence bands are expected to be occupied for $\langle 100 \rangle$, $\langle 110 \rangle$ and $\langle 111 \rangle$ respectively, thus result in the least interband phonon scattering rate for $\langle 111 \rangle$. The "camel back" band (red) is closest to top of valence band for $\langle 100 \rangle$, which may have effects on photonic characteristics of strained $\langle 100 \rangle$ nanowires. ©2010, American Institute of Physics.

2-6). In addition, the X bulk-valley at [$\pm$0.83, 0, 0] ($2\pi/a_0$) of the 3D BZ folds back into the half size 1D BZ, appearing at $\pm 0.34\pi/a_0$ of the 1D SiNW BZ [47]. The resulting fourfold Γ valley has larger confinement mass and smaller transport mass along the channel. For the off-Γ valleys, a smaller confinement mass and a larger transport mass are observed. As in the 2DOF UTB case, increased quantum confinement by the gate voltage-induced band-bending or reduced silicon thickness shifts the band energies in the conduction and valence bands. Consequently, higher confinement results in increased band gap and band splitting.

Likewise, spatially confined $\langle 110 \rangle$ channel silicon nanowires also have a CB minimum at the Γ point of the BZ. In contrast to the $\langle 100 \rangle$ channel device, only two bulk CB ellipsoids along the confined $\langle 001 \rangle$ and $\langle 00\text{-}1 \rangle$ directions, project onto the origin of the 1D reciprocal space, while four bulk CB ellipsoids appear at a non-zero k value. $\langle 110 \rangle$ SiNWs are reported to have smaller transport mass at the Γ and off-Γ valleys

compared to those of ⟨100⟩ SiNWs of same thickness [41, 48]. In addition, a larger valley splitting for the ⟨110⟩ NW triggers repopulation of carriers to the ground valleys and results in a smaller average electron conductivity effective mass compared to the ⟨100⟩ NW [41, 49, 50]. Therefore, a higher on-current is expected for a ⟨110⟩ channel SiNW compared to a ⟨100⟩ channel SiNW [51].

For higher index silicon nanowires (⟨111⟩, ⟨112⟩, etc. channel), a direct band gap is unlikely to be observed [48]. In these devices, none of the bulk originated conduction band ellipsoids reside in a direction normal to one of the confinement planes of the device, hence, projection of bulk valley(s) onto the origin of 1D BZ is not expected. Instead, these ellipsoids appear at non-zero values of the 1D wavevector [48]. Nonetheless, a direct band gap characteristic has been simulated in band calculations of ⟨111⟩ nanowires thinner than $\sim 2nm$ [41, 52, 53]. The direct band gap predicted in ultra-thin ⟨111⟩ silicon nanowires is postulated to be the result of interaction between two-sets of three-fold degenerate CB valleys situated at $\pm k_{CB_{min}}$, about the Γ point. As the thickness of the nanowire decreases, a flatter dispersion relation is seen on both sides of the Γ point, corresponding to the increase in the effective mass near the CB edge, which eventually results in a mixing the two sets of three-fold CB valleys to form a minima at the middle, which is the Γ point of the BZ [53]. Possible evidence of the direct band gap property of silicon nanowires includes a measured increase in photoluminescence (PL) intensity and efficiency and a blueshift in photon wavelength (modeled by an increase in silicon direct band gap) with decreasing nanowire diameter [39, 40, 54–56].

The effect of spatial confinement on the valence bands may be qualitatively investigated by employing a semi-analytical construction from the 3DOF bulk E-k dispersion relation [57]. In this method, holes in spatially confined nanowires are treated as “particles in a box” in the confinement directions. In order to approximate the 1DOF nanowire valence band structure, spatial confinement effects are incorporated into the

3DOF bulk silicon band structure by modeling the two confinement directions as an infinite 2D potential well. This results in the quantization of the two wavevectors along the confinement plane, with quantization intervals of $k_{1,2} = n_{1,2}\pi/L_{1,2}$, where n is an integer, and L is the thickness of the nanowire along the direction of the quantized wavevector. Then, the quantized values of the two confinement wavevectors are used to sample energy values corresponding to the continuous wavevector in the 3DOF bulk valence band dispersion. Hence, the 3D bulk valence band relation,$E(k_x, k_y, k_z)$, becomes $E(n_1\pi/L_1, n_2\pi/L_2, k_z)$ which results in individual $E_{n_1,n_2}(k_z)$ subbands for a given 1DOF nanowire cross section of L_1 by L_2. Although nonlinear effects of subband mixing and warping cannot be captured with this method, knowledge of the linear combinations of n_1, n_2 subbands can provide important insights on the valence band of silicon nanowires, without the computational costs of TB or DFT methods. Neophytou *et al.* also showed that the conduction band of 1DOF silicon nanowires can be related to the bulk silicon dispersion relations through the same semi-analytical approach [51] and suggested similar applications to estimation of extremely confined 2DOF UTB silicon band structure [51, 57].

The dispersion relations for ⟨100⟩, ⟨110⟩ and ⟨111⟩ silicon nanowire valence bands differ from each other in terms of the energy separation between the top valence bands and their curvatures (Figure 2-7). For ⟨100⟩ SiNWs, spatial confinement induced splitting of the top two valence bands is not as pronounced as in the ⟨110⟩ and ⟨111⟩ channel devices. Unlike the ⟨100⟩ SiNW VB, the light hole (LH)-like top valence band of the ⟨110⟩ SiNW resides at significantly lower hole energy than the second valence band. In addition, the ⟨111⟩ silicon nanowire valence top LH-like VB is almost fully isolated from the heavy hole (HH)-like band by several $k_B T$. Moreover, the top VB effective mass is much lighter for ⟨111⟩ and ⟨110⟩ than ⟨100⟩ SiNW. Therefore, a larger hole current is expected for ⟨111⟩ and ⟨110⟩ compared to the ⟨100⟩ SiNW [57, 58]. Furthermore, the second valence band of the ⟨100⟩ silicon nanowire has a shape of a “camel back” with

a negative effective mass (opposite band curvature than conventional positive mass dispersion) at the Γ point and two peaks at non-zero ±k values.

For gated silicon nanowires, the conduction/valence bands and electron/hole mobilities are expected to be modified through the combined spatial and electrostatic confinement effects on the carrier effective mass and scattering rates. Depending on the thickness of the nanowire (strength of the spatial confinement), the gate bias may still be effective in further splitting of the bands/subbands, resulting in for example greater isolation of the lighter mass top band [53]. Moreover, the applied electric field at different orientation surface gates and the anisotropic nature of the VB lead to preferred localization of carriers within the cross section of the device [53]. Along the confinement directions of larger mass, mobile carriers tend to reside closer to the surface. This is similar to the dependence of the carrier centroid on the out-of-plane or confinement effective masses in 2DOF MOSFETs [2]. In the 1DOF nanowire case, multiple gate orientations introduce different scattering rates governed by the roughness of the particular surface orientation and the proximity of the carriers to the gate.

2.3 Summary

Based on the carrier confinement type (electrostatic and spatial) and the degrees of freedom of the mobile carriers (3DOF, 2DOF, and 1DOF), semiconductor devices are classified in terms of their bandstructure dependent electrical properties. From the electrostatics perspective, as the devices evolve from bulk to spatially confined structures, the surface-inversion carriers spread inside the body of the crystal, leading to volume-inversion. For the quantum effects, it is shown that carrier degeneracy, electron and hole effective masses, conduction and valence band density-of-states, scattering rates and optical properties are modified with the type and the dimensionality of the confinement.

CHAPTER 3
BACKGROUND ON TRAP CHARACTERIZATION METHODS

3.1 Motivation

In order to understand the origins of the AlGaN/GaN degradation mechanisms, the nature of the electrically active traps and trap generation must be known. Due to the material challenges in GaN based devices, a significant number of point and structural defects exist in addition to intentional/unintentional impurities [59, 60]. The point defects are defined as a vacancy of a Ga (V_{Ga}) or N (V_N) atom in the lattice or a substitutional atom at these locations, as well as an interstitial atom between a Ga and a N in the unit cell. In addition to these fundamental point defects, their combinations may exist in a single GaN unit cell. For example, $V_{Ga}O_N$ is a complex point defect for which the Ga atom is missing in the unit cell and an O atom has replaced a neighboring N site. $V_{Ga}O_N$ complex can have different ion charges and is a candidate for the yellow and green band luminescence from GaN crystals [59–61]. In addition to the point defects, structural defects such as plane stacking faults, edge dislocations, threading dislocations, etc. may also adversely affect the device performance and long term reliability of AlGaN/GaN HEMTs. Further defects that contribute to device degradation are the surface states on the AlGaN barrier, and defects in the AlGaN and at the interface of the gate metal and the barrier. Consequently, the characterization of electrically active traps and understanding the fundamentals of trap generation is crucial for the realization of reliable GaN based HEMT devices.

In this regard, it is necessary to study the advantages and disadvantages of the commonly accepted trap characterization methods in terms of their measurement capabilities and experimental challenges.

3.2 Overview of Trap Characterization Methods

3.2.1 Pulsed I-V

The pulsed I-V method is based on measuring the pulsed drain current while simultaneously changing the gate voltage from the depletion condition to inversion condition and back to depletion faster than the trapping time of the mobile charge carriers in the dielectric layer [62]. In contrast to pulsed bias conditions, static DC drain current measurements suffer from an increase in the threshold voltage due to the trapped inversion layer electrons in the dielectric layer. This in return leads to a decrease in the drain current. Pulsed I-V measurements on a high-κ silicon MOSFET is shown in Figure 3-1. By adjusting the pulse width longer than the trapping time, a decrease in the drain current is observed because of the fast transient charging effects (right). Due to the decrease in the drain current during the long pulse, a hysteresis voltage difference between the rise and fall portions of the gate pulse is observed (left). The increase in the threshold voltage between the rise and fall I_D-V_G is used to extract the interface and dielectric trap densities in silicon MOSFETs.

For AlGaN/GaN HEMTs, pulsed gate voltage has been used to analyze the traps in the Ni/AlGaN interface, AlGaN barrier, and the surface states at the ungated region between the gate and the drain [63, 64]. Unlike Si MOSFETs, a lower drain current than the static DC measurements is observed for pulsed gate AlGaN/GaN HEMTs. It is shown that the dispersion between the static DC and pulsed I-V drain current decreases with increasing pulse width. This is explained by the time-dependent contribution of emitted electrons from the Ni/AlGaN and AlGaN barrier traps to the 2DEG, as well as by overcoming the time-dependent gate lag caused by charged surface states. By gradually increasing the pulse width, traps with emission time shorter than the pulse width are released, increasing the drain current. The ratio of the trap density for the corresponding emission time to the 2DEG is found by the changes in the drain current.

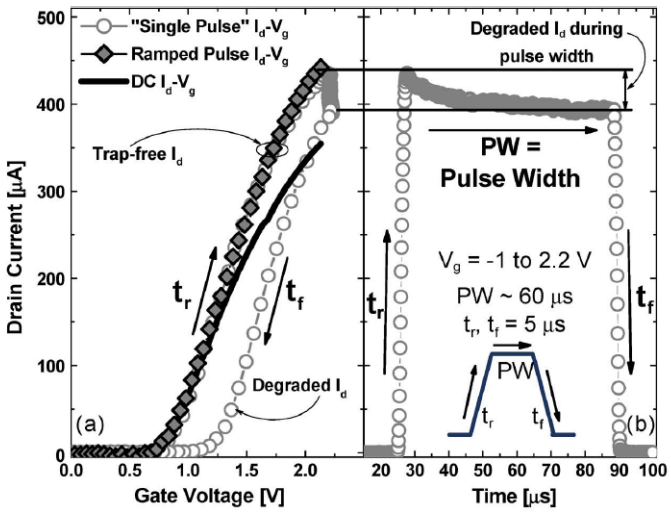


Figure 3-1. The pulsed I-V method for high-κ silicon MOSFET is shown from [62]. The gate voltage (left) and time (right) dependent drain current is plotted.

3.2.2 Conductance Dispersion

The frequency and bias dependent conductance dispersion is a well established interface trap characterization method for both silicon MOSFET and AlGaN/GaN HEMT devices. Originally developed for silicon MOSCAPs [65], the effects of interface traps are modeled as an additional capacitance (C_{it}) due to the presence of trapped charges and as a resistance (R_{it}) for the energy loss during the capture-emission processes. The equivalent circuit model of the oxide capacitance (C_{ox}), semiconductor capacitance (C_S), interface trap capacitance, and capture/emission loss is shown in Figure 3-2a. For a continuum of levels of trap states of density D_{it}, the elements of the equivalent circuit are given by [66],

$$C_P = C_s + \frac{C_{it}}{\omega\tau_{it}\,tan(\omega\tau_{it})} \tag{3–1}$$

$$\frac{G_P}{\omega} = \frac{qD_{it}}{2\omega\tau_{it}} ln[1 + (\omega\tau_{it})^2], \tag{3–2}$$

where $C_{it} = q^2 D_{it}$, ω is the angular frequency of capacitance measurement, and $\tau_{it} = R_{it}C_{it}$ is the interface trap time constant. By measuring capacitance at different frequencies, the capacitance and conductance dispersion curves are obtained. From these, the interface trap density is extracted from the maximum of the G_P/ω and the

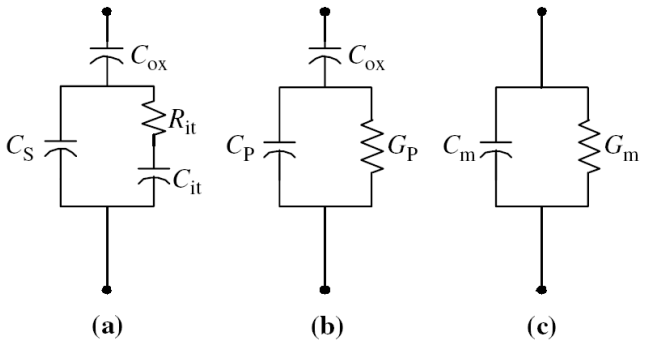


Figure 3-2. (a) Silicon MOSCAP capacitance is modeled with the oxide capacitance, semiconductor capacitance, interface trap capacitance and capture/emission loss. (b) Equivalent circuit with interface trap conductance. (c) Circuit of the measured capacitance and parallel conductance.

interface trap time constant is found from the angular frequency for which the peak conductance is measured.

The conductance dispersion method has been used to extract the trap density, time constant, and activation energy for GaN based HEMT devices [67–70]. The differences between silicon MOSCAPs and GaN based HEMT devices are represented as additional elements to the equivalent circuits for the AlGaN surface trap states, series resistance and multiple interface and bulk GaN traps. By modifying the bias conditions, depletion or enhancement of the 2DEG is achieved for extraction of traps from different spatial locations of the HEMT structure (i.e interface or bulk). Activation energies and capture cross sections of the modeled traps are obtained by varying the temperature during the dispersion experiments. However with an increasing number of resistive and capacitive elements in the model circuits, the accuracy decreases due to the fact that capacitance measurements can only provide the total effective capacitance and conductance. This leads to an equation set with too many unknowns and two measured data. The individual trap parameters are approximated by mathematical fitting models which may be either too complicated or inaccurate.

3.2.3 Subthreshold Swing

The subthreshold drain current measurement for trap characterization provides an average interface trap density for MOSFETs [66]. The subthreshold swing is modified by the amount of charge trapped at the interface states and is given by,

$$S = \frac{KT}{q} ln(10) \frac{C_b + C_{it}}{C_{ox}}, \tag{3–3}$$

where C_b is the space-charge region capacitance. In order to extract the interface trap capacitance (i.e density), the exact values of the oxide and depletion capacitance is required. Hence, the subthreshold swing method is generally employed as a simple differential trap characterization method for devices before and after degradation. However by varying the temperature, Chung *et al.* has demonstrated that the subthreshold swing method can provide a quick estimate of the average interface trap density for AlGaN/GaN HEMTs [71].

3.2.4 Current Transients

Investigation of trapping and detrapping behaviors through current transients is one of the earliest methods of trap parameter extraction from semiconductor devices [72, 73]. Although the method was developed by C. T. Sah almost 40 years ago, it is still currently used as a fast trap characterization technique for AlGaN/GaN HEMT devices [25]. Bulk GaN traps, AlGaN barrier traps, and AlGaN surface states may trap energetic channel electrons which then lead to partial depletion of the channel, and hence to a current collapse [22, 23]. Joh and del Alamo examined the trapping and detrapping behavior of the current collapse and its recovery using current transients [25]. Both measured transients were fitted to a linear combination of exponentials with different time constants. By exploring the temperature dependent changes in the fitted time constants, trap activation energies were estimated using Arrhenius plots. Additionally, by changing the device operation condition, information on the spatial location of the trap is estimated.

Current transients is a simple and fast trap characterization technique, which is useful for semi-quantitative measurements. The disadvantages of this method are the requirement of a cryogenic temperature range and the inability to extract trap density.

3.2.5 Deep Level Transient Spectroscopy (DLTS)

The fundamentals for the deep level transient spectroscopy method were developed by C. T. Sah and his students in the early 70's [72, 74, 75]. Later, Lang has implemented the method with a double boxcar integrator [76]. In DLTS, the transient capacitance, charge, or current from a reverse bias pulsed pn junction or a Schottky diode is monitored at a desired pair of sampling times (t_1 and t_2) under a slowly varying temperature. When plotted with respect to the measurement temperature (T), the resulting differential capacitance (or charge,current) of the transient at t_1 and t_2 forms a local maximum for each trap. The emission time (τ_e) is fixed for each chosen pair of t_1 and t_2, and determines the temperature for which the peak is observed. At each local maxima, the density of the trap is obtained with the help of the value of the transient peak. By changing t_1 and t_2, and performing a new temperature sweep while measuring the transients, a new differential peak is observed at a different temperature. Hence, a new set of τ_e and T is found for each trap. Consequently, the $\ln(\tau_e T^2)$ vs $1/T$ (Arrhenius) plots are obtained for individual traps. Using the Arrhenius plots, the trap activation energy and capture cross sections are found.

The DLTS based approaches have been used for trap characterization in GaN based material systems [77–81]. Okino *et al.* have investigated the trap characteristics of a metal/SiN/AlGaN/GaN HEMT through current-DLTS measurements while sweeping the temperature [81]. Faqir *et al.* have also employed the temperature dependent current-DLTS method for the analysis of trap related effects in AlGaN/GaN HEMTs [80]. It was also shown that a combination of photon energy (deep level optical spectroscopy - DLOS) and temperature sweep can be used in a complementary manner to scan different parts of the GaN energy gap during capacitance or current measurements

[77–79]. In [79], the gate and drain voltages were dynamically controlled to achieve a constant drain current at different operation regimes during DLOS and DLTS experiments. By monitoring the changes in the feedback response of the gate and drain voltages, the distribution of traps under the gate and near the drain access region was estimated [79].

DLTS measurements are powerful tools to investigate traps in AlGaN/GaN devices however, the requirement of a cryogenic to elevated temperature range with optical illumination where needed results in a very challenging experimental setup when combined with mechanical wafer bending experiments.

3.2.6 Photoionization Spectroscopy and Luminescence

The density, energy, and trapping time of subbandgap traps in AlGaN/GaN HEMT devices can be determined by photoionizing the trapped electrons with a varying optical excitation wavelength and measuring the changes in the time dependent drain current. Photoionization spectroscopy has been employed for the investigation of traps via the reversal of drain current collapse in GaN devices due to the captured hot electrons in slow bulk GaN traps and impurity sites [82–86]. For a trap state with an activation energy of E_A, optical illumination from an incident photon with an energy equal to or greater than E_A results in the emission of the trapped electron to the GaN conduction band. With the assumption that the time required to traverse the HEMT channel is shorter than the trap capture time, the trap state at E_A below the conduction band is continuously emptied to a non-equilibrium steady-state, under a sufficiently high density of incident photon flux. Hence, the electrons residing at multiple traps with activation energy lower than the incident photon energy contribute to the 2DEG and increase the drain current during the optical illumination. Assuming that the carrier mobility remains unchanged, the ratio of the total excited trap density to the 2DEG is obtained by comparing the increased drain current under illumination to the dark current. Using a range of different subbandgap light wavelengths, the energy distribution of the areal trap

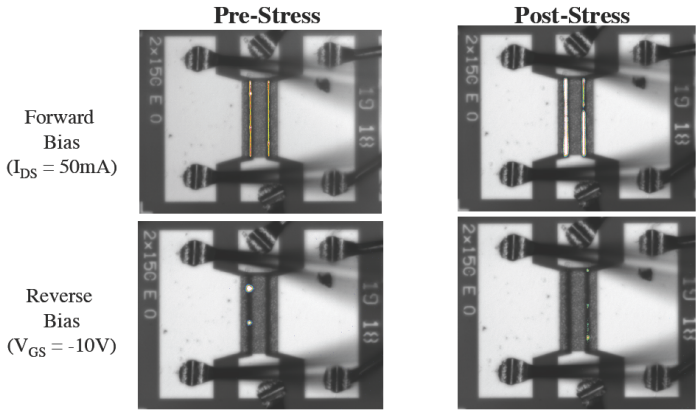


Figure 3-3. The photoemission is shown for pre-stress and post-stress conditions under forward and reverse bias [93].

density may be extracted. After the light source is switched off, the emitted electrons are captured by the emptied trap states, leading to a decaying transient in the drain current. By fitting exponential functions to the decay transient, trap capture times can be extracted.

The luminescence properties of the GaN point defects, structural defects, and impurities have been studied thoroughly in the literature [59–61, 87–93]. The previous investigations of the defects responsible for red, yellow, green, blue and near bandgap optical emission bands provides invaluable information on the nature of the traps obtained from optical excitation experiments. Moreover, since the traps are extracted by optical excitation, the spatial location of the traps can be determined through the photoluminescence, cathodoluminescence and electroluminescence response of the device for the given radiative trap energy (Figure 3-3).

3.3 Summary

Based on the measurement capabilities and experimental challenges of trap characterization methods, Table 3-1 summarizes a comparison of different techniques. The trap activation energy (E_A), corresponding energy distribution of areal trap density (D_{it}), and trapping time constant (τ) are required to quantify the trapping behavior in AlGaN/GaN HEMTs. In this regard, conductance dispersion, DLTS and photoionization spectroscopy can provide most of the required parameters. However as we have

Table 3-1. Summary of different trap characterization methods in terms of capabilities of density (D_{it}), time constant (τ), and activation energy (E_A) extraction, and experimental challenges.

Method	Can Measure	Cannot Measure	Experimental Challenges
Pulsed I-V	D_{it}, τ	E_A	nanosecond RF probing
Conductance Dispersion	D_{it}, τ, E_A	N/A	cryogenic to elevated T range
Subthreshold Slope	D_{it}^{avg}	E_A,τ	elevated T range
Current Transients	τ, E_A	D_{it}	cryogenic to elevated T range
DLTS	D_{it}, τ, E_A	N/A	cryogenic to elevated T range + optical illumination
Photoionization Spectroscopy	D_{it}, τ, E_A	N/A	optical illumination

discussed, the conductance dispersion method suffers from the limited number of measured quantities (total conductance and capacitance) compared to the number of capacitive and resistive trap elements to be modeled. Also, the requirement of cryogenic temperature range while conducting external wafer bending increases the complexity of the experimental setup significantly. In addition, the DLTS technique requires both cryogenic temperature range and optical illumination during wafer bending, which further increases the complexity of the experimental setup. On the other hand, photoionization spectroscopy can be carried out by illuminating the sample in the wafer bending jig, without the need for a cryogenic chamber. Thus, photoionization spectroscopy is an experimentally feasible method for the characterization of strain effects on AlGaN/GaN HEMT trap generation.

CHAPTER 4
INSIGHTS ON THE (100) AND (110) DOUBLE-GATE FINFET ELECTRON TRANSPORT AND THE STRAIN RESPONSE OF P-TYPE TRI-GATE FINFETS

4.1 Physical Insights on Comparable Electron Transport in (100) and (110) Double-gate Fin Field-effect Transistors

4.1.1 Motivation

As the channel length of complementary metal-oxide-semiconductor (CMOS) transistors is scaled down to sub-100nm length range, detrimental short channel effects (SCE) have mitigated device performance. To overcome these effects, non-planar silicon channels with multiple gates are introduced to restore the gate control on short-channel on/off resistance. In fact, Intel has recently announced their 22nm baseline comprised of Tri-Gate FinFETs[7].

In addition to better SCE control, the non-planar transistor approach provides both (100) and (110) sidewall transport on the same conventional (100) silicon wafer, avoiding more expensive solutions such as Hybrid Orientation Technology (HOT)[94]. Double-gate FinFETs (DG-FinFETs) fabricated on (100) wafers have (110)-surface/$\langle 110 \rangle$-channel or (100)-surface/$\langle 100 \rangle$-channel configurations. For conventional planar metal-oxide-semiconductor field effect transistors (MOSFETs) fabricated on separate wafers, (100) is more favorable for electron conduction, whereas (110) is preferable for hole transport.

Contrarily, we have experimentally observed comparable electron mobility in (110) and (100) sidewall DG-FinFETs[95], while (110) is still favorable for hole transport over (100). Therefore, a physics based investigation is very timely for selection of

---

Section 4.1 is reprinted with permission from Mehmet O. Baykan, Chadwin D. Young, Kerem Akarvardar, Prashant Majhi, Chris Hobbs, Paul Kirsch, Raj Jammy, Scott E. Thompson, and Toshikazu Nishida, “Physical insights on comparable electron transport in (100) and (110) double-gate fin field-effect transistors” Appl. Phys. Lett. 100, 123502 (2012), DOI:10.1063/1.3696038. Copyright 2012, American Institute of Physics.

the next generation CMOS integration scheme on (110) DG-FinFET platform. In this section, we investigate one of the possible underlying physical reasons for comparable electron mobility seen in (100)/⟨100⟩ and (110)/⟨110⟩ FinFETs, from a perspective of non-parabolic dispersion relation of transverse electron mass (m_t) in the ⟨110⟩ direction.

4.1.2 Experimental

Silicon-on-insulator (SOI) DG-FinFETs were fabricated with channels formed on (100) and (110) planes. Neutral stress (100) intrinsic ($N_D \approx 2x10^{15} cm^{-3}$) SOI substrates were patterned using 193 nm lithography to produce arrays of 200 fins with 10 μm channel length. Hafnium based high-k (∼2 nm) and mid-gap metal (∼10 nm) gate stack was formed. The resulting fins have cross section dimensions of 40 nm height and 20 nm width. Split capacitance-voltage (*C*-*V*) measurements were conducted to extract effective electron mobility ($\mu_{n,eff}$).

Figure 4-1 shows the extracted $\mu_{n,eff}$ vs. inversion electron density (N_{inv}) from a total of 25 sites on two different 200mm and 300mm wafers manufactured at different facilities. Unlike the conventional planar n-type MOSFETs fabricated on separate (100) and (110) wafers, no degradation in mobility is observed as the sidewall is changed from (100) to (110). Previously, it was suggested that high-κ/metal-gate induced stress could be the reason for the observed anomaly [96]. However, we have recently shown that mobility extracted from SiO_2/poly-Si gate FinFETs also demonstrate comparable electron mobility for (100) and (110) sidewalls[95, 97]. Similarly, we have observed the same phenomenon from fins with varying thickness from 20nm to 45nm[95], as well as high dose P or B doping[97]. Hence, the reason for the comparable electron mobility in (100) and (110) sidewall fins could not be explained by any combination of the metal-gate induced stress, volume-inversion, and undoped body effects.

To understand the underlying physical mechanisms, we have revisited the conduction band of (100) and (110) FinFETs. The inset of Figure 4-1 shows the conduction band (CB) ellipsoids and their conductivity effective mass (m^*_{cond}) for

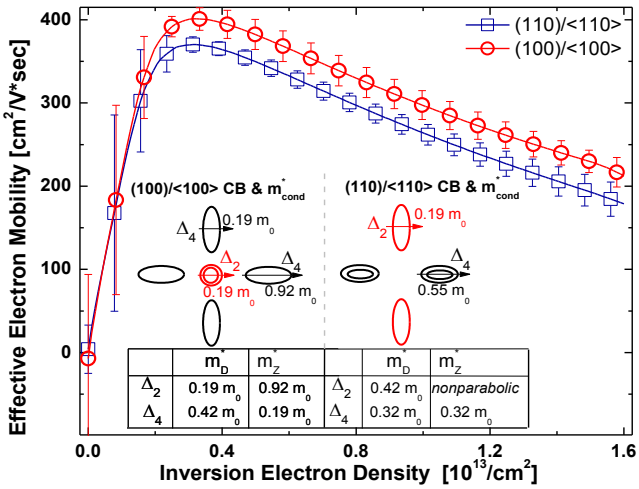


Figure 4-1. Effective electron mobility is extracted from (100) and (110) sidewall FinFETs using split-CV method, and shown as a function of N_{inv}. Inset shows the conduction band (CB) ellipsoids and their conductivity effective masses for (100) and (110) inversion layer. In addition, the density-of-states (m_D^*) and confinement mass (m_Z^*) values are also tabulated. ©2012, American Institute of Physics.

both (100) and (110) inversion layers. The Δ_2 valley m_{cond}^* is $0.19m_0$ for both sidewall orientations. On the other hand, Δ_4 valleys of (100) have two different transport masses with equal occupancy, resulting in average $m_{cond}^* = 0.315m_0$. Similarly, the m_{cond}^* of Δ_4 valleys in (110) is found to be equal to $0.55m_0$. The table shown in the Figure 4-1 inset lists the density-of-states (DoS) and confinement effective mass values (m_D^* and m_z^* respectively). Although generally taken as parabolic, with $m_t = 0.19m_0$, only the m_z^* of (110) Δ_2 is non-parabolic and has significant effects on electron transport in (110). The effects of non-parabolic confinement mass ($m_{z,NP}^*$) for (110) have been considered by Uchida[98], Silvestri[99] and Shimizu[100]. Yet, none of these efforts could explain our experimental results shown in Fig.4-1.

In order to obtain the average conductivity effective mass ($m_{cond,avg}^*$) for both (100) and (110) FinFETs, a self-consistent single electron effective mass Schrödinger-Poisson DG-FinFET simulator is used. Additionally, to investigate the effects of $m_{z,NP}^*$, we have employed a nearest-neighbor $sp^3d^5s^*$ tight-binding (TB) formalism[101, 102] to obtain the *E-k* relation of 20nm thick (110) and (100) FinFET supercells. It is verified that only

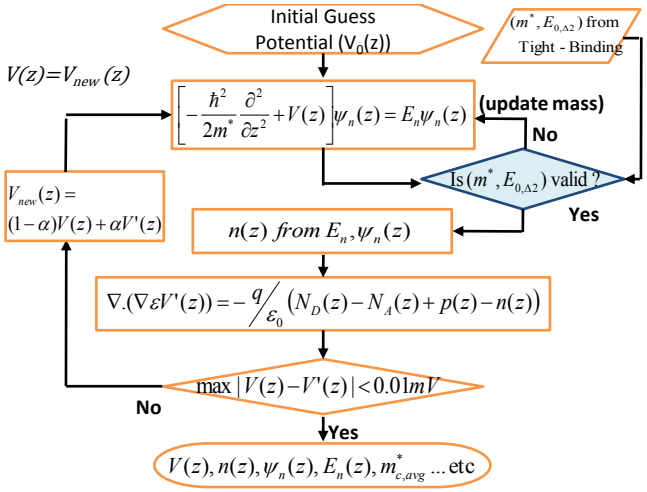


Figure 4-2. Standard Schrödinger-Poisson self-consistent simulation procedure is modified to take non-parabolic confinement mass ($m^*_{z,NP}$) of Δ_2 valley on (110) into account. The inner loop ensures that the correct $m^*_{z,NP}$/quantization energy pair converges. ©2012, American Institute of Physics.

Δ_2 of (110) FinFET has non-parabolic m^*_z. Using the TB results, we obtained subband quantization energy dependent $m^*_{z,NP}$. Figure 4-2 shows the procedure followed during the self-consistent Schrödinger-Poisson simulation. For the (110) FinFET, we have modified the procedure with an inner control loop to select the correct ($m^*_{z,NP}/E_{0,\Delta_2}$) pair obtained from the TB calculations.

4.1.3 Results and Discussion

Conventionally, a parabolic fit is used to approximate (110) Δ_2 confinement mass as $m_t = 0.19m_0$. Since the valley subband energies are inversely related to m^*_z, and the electrons preferably occupy lower energy states, the parabolic confinement mass approach (PCM) assumes that the majority of electrons occupy Δ_2 for (100) and Δ_4 for (110). This approach inherently leads to heavier $m^*_{cond,avg}$ for (110), and hence, much lower $\mu_{n,eff}$ compared to (100) sidewall DG-FinFETs.

However, the *E-k* dispersion of m_t in the $\langle 110 \rangle$ direction is strongly non-parabolic, which results in a non-parabolic confinement mass for the (110) Δ_2 valley. As shown in Figure 4-3-a, the *E-k* curve deviates from the parabolic band as the energy increases.

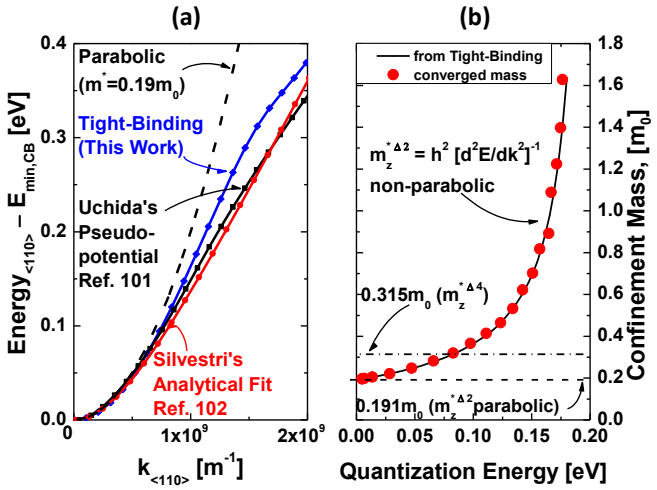


Figure 4-3. (a) Energy-k dispersion relation of m_t in $\langle 110 \rangle$ direction is shown for parabolic and non-parabolic band models. (b) The solid line represents the $m^*_{z,NP}$ of $\Delta_2^{(110)}$ is obtained from the inverse second derivative of the tight-binding *E-k* relation. Circles are the converged ($m^*_{z,NP}/E_{0,\Delta_2}$) pairs in the inner loop of Fig.4-2. ©2012, American Institute of Physics.

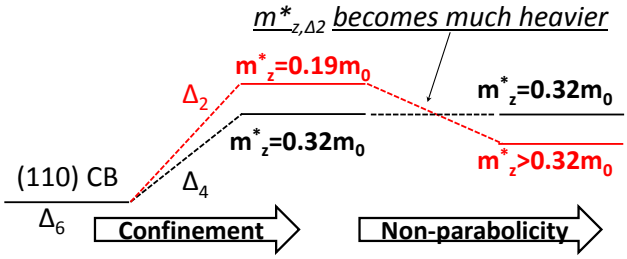


Figure 4-4. (110) CB valley energy and degeneracies are shown. Under PCM, confinement splits valleys and Δ_4 becomes ground valley due to the heavier m^*_z. When non-parabolicity is included, the m^*_z of Δ_2 becomes much heavier and switches the valleys over. ©2012, American Institute of Physics.

In fact, the $m^*_{z,NP}$ quickly becomes heavier than m_t with increasing quantization energy (Figure 4-3-b). Seen as the circles on Figure 4-3-b are the converged ($m^*_{z,NP}/E_{0,\Delta_2}$) pairs in the inner control loop of the simulator. Consequently, a very good agreement between TB and Schrödinger-Poisson domains has been achieved.

Heavier non-parabolic confinement mass results in lower (110) Δ_2 energy for a given N_{inv}, thus increasing its occupancy. With the increased occupancy of the Δ_2 valley, the percentage of electrons with the lighter m_t conductivity effective mass increases.

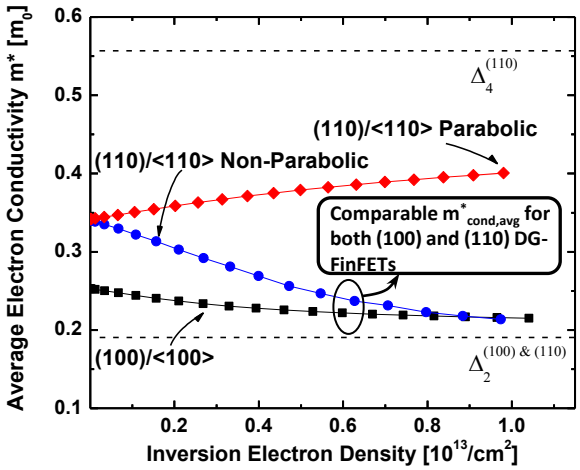


Figure 4-5. Inversion electron density dependent $m^*_{cond,avg}$ is shown for (100) and (110) for both parabolic and non-parabolic confinement mass methods. Under non-parabolic model, the $m^*_{cond,avg}$ for both (100) and (110) are similar, which suggests the comparable electron mobility shown in Fig.4-1. ©2012, American Institute of Physics.

Therefore, the (110) average m^*_{cond} decreases with stronger non-parabolic effects on the confinement mass. Figure 4-4 demonstrates the effect of non-parabolic confinement mass for (110) CB. Previous reports on the $m^*_{z,NP}$ for (110) have indeed showed higher Δ_2 occupancy[98–100]. However, these prior efforts resulted in higher Δ_2 valley energy than Δ_4, indicating $m_t < m^*_{z,NP} < 0.32m_0$, even at high N_{inv}. On the contrary, our non-parabolic confinement mass calculations suggest that m^*_z of Δ_2 equals to Δ_4 at a quantization energy of about $83meV$ or $N_{inv} \sim 1.6 \cdot 10^{12}cm^{-2}$. Furthermore, at higher N_{inv}, Δ_2 and Δ_4 switches over and Δ_2 becomes the most occupied valley.

With the inclusion of the non-parabolicity in confinement mass, our DG-FinFET simulation results qualitatively explain the experimentally observed similarity in (100) and (110) $\mu_{n,eff}$. Shown in Figure 4-5 are the inversion electron density dependent $m^*_{cond,avg}$ for (100) and (110) FinFETs. Using the non-parabolic mass for (110) FinFET, the Δ_2 valley becomes the ground valley and thus the $m^*_{cond,avg}$ of (110) asymptotically approaches to $0.19m_0$. Likewise, the (100) FinFET Δ_2 valley becomes the dominant transport valley with increased N_{inv}, and the average mass decreases towards $0.19m_0$.

In fact, at high N_{inv}, the (110) mass is lower than (100) due to larger DoS at Δ_2 valley of (110) compared to (100). Yet, higher DoS of (110) would increase the scattering rates with respect to (100), counteracting the lighter mass. Consequently, (100) and (110) DG-FinFET electron mobility are expected to be comparable, qualitatively explaining the experimental data shown in Figure 4-1.

4.1.4 Conclusion

By including the non-parabolic effects on the confinement mass of Δ_2 valley of (110) conduction band, we have explained a possible physical mechanism governing the comparable electron mobility in (100) and (110) DG-FinFETs. Although the shortcomings in explaining planar MOSFET behavior are yet to be investigated, it is shown that $m^*_{cond,avg}$ for (100)/$\langle 100 \rangle$ and (110)/$\langle 110 \rangle$ FinFETs becomes similar with the inclusion of non-parabolicity of confinement mass. Consequently, without degradation in electron transport, the (110) sidewall FinFET is a cheaper alternative to HOT for high performance device integration at and beyond the 22nm node.

4.2 Strain Response of p-type Tri-Gate FinFETs

Following the efforts of understanding the electron transport in double-gate FinFETs in the previous section, we continue with the investigation of the strain enhancement of p-type FinFETs, particularly to achieve balanced pull-up and pull-down elements in the CMOS implementation of this non-planar technology. Due to the non-planar structure of the tri-gate device, carriers transport through different surface orientations such as (110) for the sidewall and (100) for the top surface. In this regard, we have conducted an experimental study of 4-point external mechanical bending of p-type tri-gate FinFETs with a fixed height (25nm) and a varying width from 25nm to 500nm (Figure 4-6). Figure 4-7 shows the fundamentals of the 4-point uniaxial wafer bending approach. It is shown that the portion of the sample between the inner rods is under a constant uniaxial stress

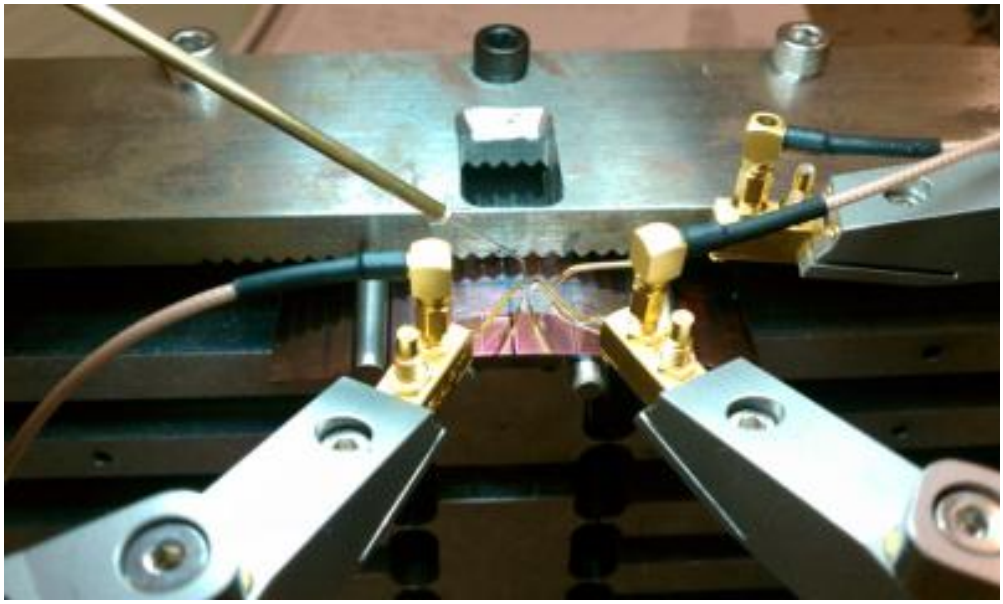


Figure 4-6. Tri-gate p-type FinFET sample cleaved from a 300mm wafer is under ∼1GPa uniaxial tension.

given by [103],

$$\sigma = \frac{Eyt}{2a(\frac{L}{2} - \frac{2a}{3})}, \tag{4–1}$$

where σ is the uniaxial stress, E is the Young's Modulus of the substrate, t is the wafer thickness, y is the vertical displacement, a is the inner rod spacing and L is the outer rod spacing. In addition to the calculated value from Eq. 4–1, the amount of applied strain is also measured by a strain gage attached to the wafer sample during the application of external mechanical bending.

The strain response of the tri-gate FinFETs is characterized by the piezoresistance (π) coefficient. The π-coefficient is defined as the normalized change in the resistance per stress, i.e.

$$\frac{\Delta R}{R}\frac{1}{\sigma} = \pi, \tag{4–2}$$

where σ is the applied mechanical stress. To measure the piezoresistance coefficients, low field drain current is measured from 11 devices with varying widths from 25nm to 500nm, under $V_{DS} = 0.05V$ and $V_{GS} = -0.5V$ to $1.5V$, up to 250MPa longitudinal

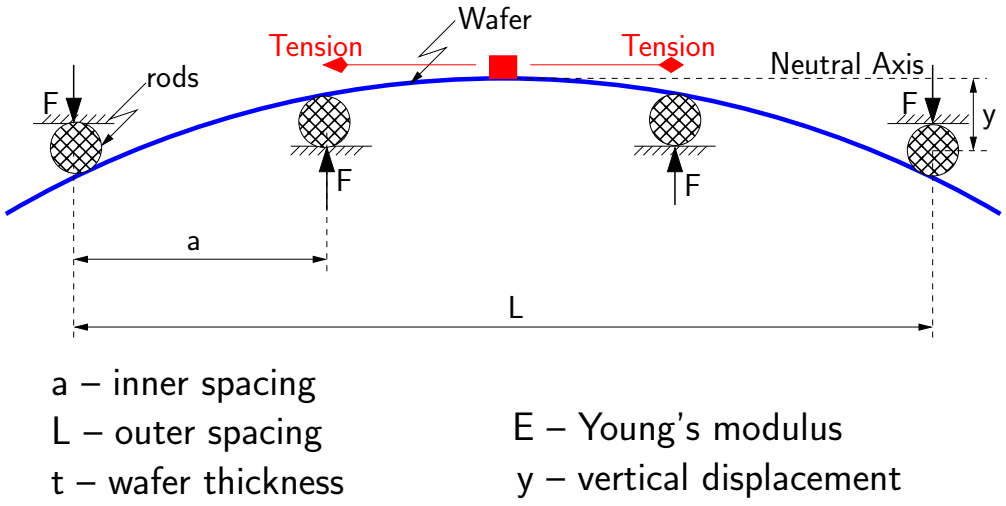


Figure 4-7. 4-point uniaxial wafer bending is illustrated.

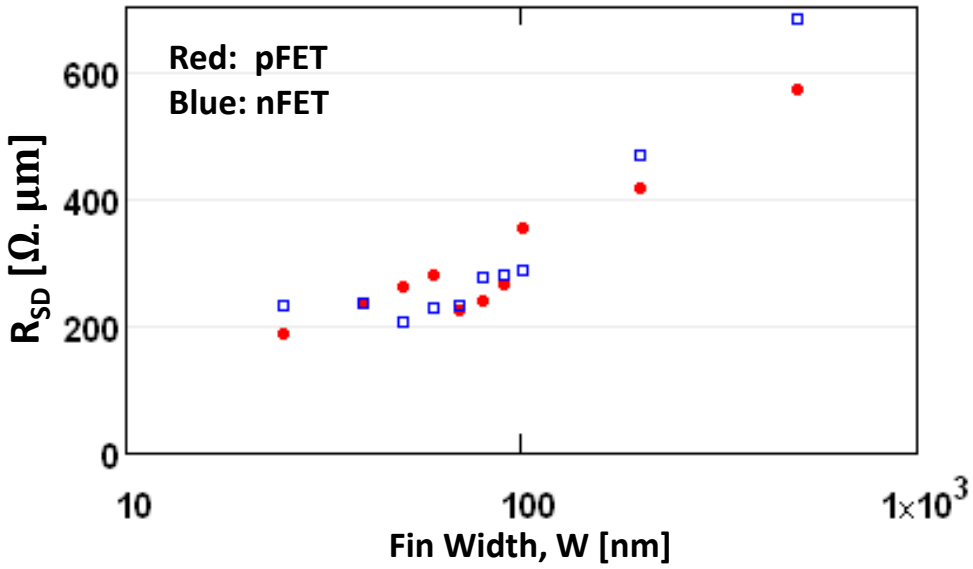


Figure 4-8. The external parasitic resistance R_{SD} extracted from n (blue) and p-type (red) tri-gate FinFETs using shift and ratio method.

uniaxial compression. The linear device current is corrected using the extracted R_{SD} contribution shown in Figure 4-8. The piezoresistance coefficients are extracted at the strong inversion limit ($V_{GS} - V_{TH} = 0.5V$) and are shown as symbols in Figure 4-9.

In order to model the experimentally observed width dependence of the π-coefficient of the tri-gate FinFETs, we have considered the bulk MOSFET π-coefficients of the top and sidewall surfaces [104]. As will be discussed in Chapter 6, volume-inversion related

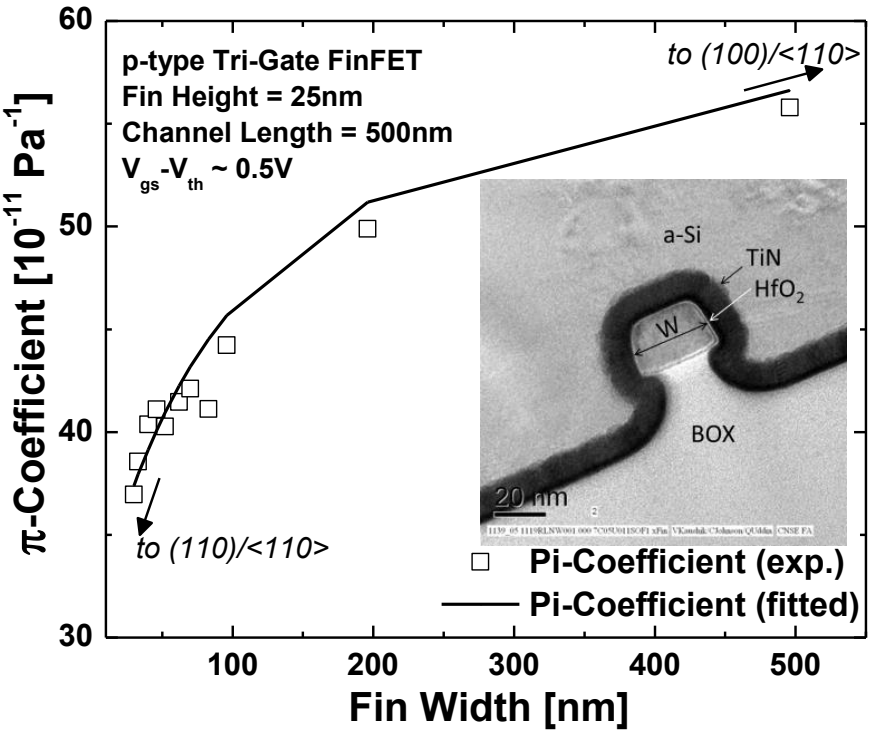


Figure 4-9. Experimentally extracted piezoresistance coefficients from p-type tri-gate FinFETs are shown (symbols). Solid line denotes the calculated π-coefficient using Eq. 4–8. Close agreement between experimental and calculated values is obtained for $\pi_{(100)} = 65 \cdot 10^{-11} Pa^{-1}$, $\pi_{(110)} = 30 \cdot 10^{-11} Pa^{-1}$ and $\mu_{(110)}/\mu_{(100)} = 2$. Inset shows the cross section TEM of a 30nm wide tri-gate device.

effects are minimal for relatively thick FinFETs of 20nm or wider, which justifies the use of bulk surface inversion MOSFET π-coefficients. Hence, the tri-gate FinFET is assumed to consist of 2 parallel channels, namely the (100) top surface and the (110) sidewall surface. For the FinFET with height H, width W and gate length L, the width of the (100) and (110) channels are equal to W and 2H respectively. Then the conductances of the top and sidewall channels are given as,

$$G_{(100)} = \frac{W \mu_{(100)} C_{ox,100} (V_G - V_{T,100})}{L} \quad (4–3)$$

$$G_{(110)} = \frac{2H \mu_{(110)} C_{ox,110} (V_G - V_{T,110})}{L} \quad (4–4)$$

$$G_{TG} = G_{(100)} + G_{(110),} \quad (4–5)$$

where μ is the surface mobility, $C_{ox}(V_G - V_T) = qN_{inv}$ denotes the inversion areal carrier density, and G_{TG} the total FinFET conductance. In terms of conductance, the π-coefficient can be defined as,

$$\frac{G^0 - G^\sigma}{G^\sigma} = \pi\sigma, \tag{4–6}$$

where G^0 and G^σ are the unstrained and strained conductances respectively. Then the tri-gate π-coefficient, π_{TG} can be written as,

$$\pi_{TG} = \frac{1}{\sigma}\frac{\dfrac{W\mu_{(100)}qN_{(100)} + 2H\mu_{(110)}qN_{(110)}}{L} - \left[\dfrac{W\mu_{(100)}qN_{(100)}}{L(1+\pi_{(100)}\sigma)} + \dfrac{2H\mu_{(110)}qN_{(110)}}{L(1+\pi_{(110)}\sigma)}\right]}{\dfrac{W\mu_{(100)}qN_{(100)}}{L(1+\pi_{(100)}\sigma)} + \dfrac{2H\mu_{(110)}qN_{(110)}}{L(1+\pi_{(110)}\sigma)}} \tag{4–7}$$

where $N_{(100)}$ and $N_{(110)}$ are the inversion carrier densities for top and sidewall surfaces, which are assumed to be similar. For low stress conditions up to 200MPa, both $\pi_{(110)}\sigma$ and $\pi_{(100)}\sigma$ terms are much smaller than 1, and then the tri-gate π-coefficient is found by,

$$\pi_{TG} = \frac{W\pi_{(100)} + 2HK\pi_{(110)}}{W + 2HK}, \tag{4–8}$$

where $K = \mu_{(110)}/\mu_{(100)}$. Using Eq. 4–8, we have plotted the calculated tri-gate π-coefficient as a function of fin width using $\pi_{(100)} = 65\cdot10^{-11}Pa^{-1}$, $\pi_{(110)} = 30\cdot10^{-11}Pa^{-1}$ and $\mu_{(110)}/\mu_{(100)} = 2$. The $\pi_{(100)}$ and $\pi_{(110)}$ values that provide the best fit to experimental data are well within the published experimental values [104]. Consequently, for non-planar structures wider than 20nm, the piezoresistance coefficients can be obtained by using a weighted conductance average of the bulk MOSFET π-coefficients of the contributing transport surfaces.

4.3 Summary

We have conducted a coupled experimental and modeling study to investigate the underlying physics of comparable electron transport in (100) and (110) double-gate FinFETs and the width dependence of the strain response of p-type tri-gate FinFETs. One possible reason for the comparable electron transport is shown to be the nonparabolic confinement effective mass of Δ_2 CB valley in (110) FinFETs, resulting in a similar

average electron conductivity effective mass with respect to (100) FinFETs. On the other hand, the strain response of p-type tri-gate FinFETs is shown to be dependent on fin width, and can be accurately modeled by using the conductivity weighted average of the top and sidewall surface bulk MOSFET piezoresistance coefficients.

CHAPTER 5
SIZE AND ORIENTATION DEPENDENT STRAIN EFFECTS ON BALLISTIC SI P-TYPE NANOWIRE FIELD EFFECT TRANSISTORS

5.1 Motivation

The continuous scaling down of planar silicon transistor dimensions for over four decades has introduced an unavoidable barrier of detrimental short-channel effects. As the devices approach the quasi-ballistic transport limit, gate control of the channel resistance has degraded notably. One solution to overcome this problem is the non-planar multiple-gate transistor design, which helps to restore the gate control over the channel resistance [105]. Due to their excellent short-channel control capability, silicon based FinFET and nanowire (NW) devices have been considered as the next possible candidates for the near future technology nodes. In fact, Intel has recently announced their 22nm baseline comprised of Tri-Gate FinFETs[7]. For devices at the scaling limit beyond the 22nm node, carriers are expected to traverse the transistor channel at near ballistic transport conditions [12]. Hence, the nanowire based device design may overcome the stronger short-channel effects in further scaled devices beyond the 22nm technology node.

In recent years, process induced external mechanical strain has become the most established device performance booster since Intel's 90nm technology [1]. Strained advanced CMOS devices have enjoyed superior performance compared to unstrained ones through strain-induced reduction of the carrier effective mass and scattering rates [2, 106]. Not surprisingly, strain is also expected to be employed as a performance booster in the future ballistic devices [11, 107]. The incorporation of mechanical strain

Reprinted with permission from Mehmet O. Baykan, Scott E. Thompson, and Toshikazu Nishida, "Size and Orientation Dependent Strain Effects on Ballistic Si p-type Nanowire Field Effect Transistors" IEEE Transactions on Nanotechnology, 2012. Copyright 2012, IEEE.

into the atomistic nanowire device simulations is of utmost importance for performance evaluation of near future device designs.

In order to model and predict future strained silicon CMOS platforms, there is a need for a detailed physics based investigation of the strain enhanced ballistic p-type nanowire transport. Prior work showed through extensive self-consistent atomic simulations that the $\langle 110 \rangle$ orientation provides the highest current for both p and n-type unstrained ballistic nanowire transistors [51, 57, 108]. To increase computational efficiency, compact electrostatic models were coupled to atomic band structure simulations to investigate ballistic nanowire electron current [109, 110]. Recently, a study of strain effects on n-type ballistic nanowire transport was published by Zhang *et al.* [111]. Therefore, merging atomistic band structure calculations with recent developments in compact ballistic transport models [112, 113] is very timely for understanding the ballistic hole current enhancement in strained p-type nanowires. From a manufacturing point of view, strain could also mitigate the effects of process-induced thickness variation for p-type ballistic nanowires [114]. Due to the dependence of spatial confinement effects on the device size and orientation, physics of the strain enhancement of ballistic p-type nanowire transistor performance needs to be evaluated and compared as a function of size and device orientation for future high-performance CMOS integration platforms.

In this chapter, we explain the physical mechanisms which govern the longitudinal uniaxial tensile and compressive strain responses of $\langle 110 \rangle$ and $\langle 100 \rangle$ Si ballistic nanowire p-type field-effect transistors of three different dimensions: $3nm \times 3nm$, $5nm \times 5nm$ and $7nm \times 7nm$.

5.2 Method

Here, we consider straight nanowire devices extending in the $\langle 110 \rangle$ direction with (001) and $(\bar{1}10)$ gate surfaces, while the rotated device channel is in the $\langle 100 \rangle$ direction with equivalent (100) gates (Figure 5-1). Conventional (100) wafers are assumed for both

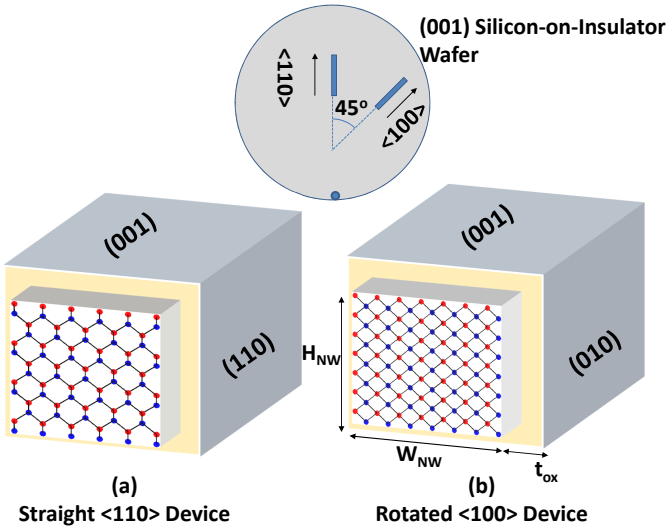


Figure 5-1. Straight $\langle 110 \rangle$ and 45° rotated $\langle 100 \rangle$ devices are shown on conventional (001) silicon-on-insulator wafer. The unit supercell is represented using "ball and stick" model of covalent bonds. ©2012, IEEE.

straight transistors in the $\langle 110 \rangle$ direction along the wafer cut and 45° rotated transistors in the $\langle 100 \rangle$ direction. Due to the quantized spacing between silicon atoms in different crystal planes, actual device dimensions are slightly different than the considered values. The actual atomistic device dimensions can be seen in Table 5-1. Finally, all types of devices considered in this work have surrounding gates with an effective oxide thickness of $1.1nm$.

Due to the atomistic nature of nanowire dimensions, nearest-neighbor $sp^3d^5s^*$ tight-binding (TB) formalism is employed for strained and unstrained band structure calculations [115] and the updated Si tight-binding parameters from Boykin's work are used [116, 117]. The nanowire unit supercell is created for given material, orientation and cross-sectional dimensions. The resulting supercell includes all the atoms in the 2-dimensional cross section of the nanowire and allows us to assume 1-dimensional plane-wave propagation in the channel length direction. Hence, the calculated band structure leads to an energy vs. 1D wavevector dispersion relation in the transport direction. The choice of $sp^3d^5s^*$ basis set with spin orbit coupling implies a $20N \times 20N$

Hamiltonian matrix, where N is the number of atoms in the 2D nanowire supercell. This choice allows us to capture the detailed band structure of these nanowires for evaluation of the strain effects on the ballistic hole transport. Due to the very high surface to volume ratio present in nanowire structures, the dangling surface bonds of the finite crystal are passivated to avoid significant amount of mid gap eigenvalues in the band structure [118].

The mechanical strain is incorporated via elastic deformation of the two-dimensional nanowire supercell. Although Leu *et al.* showed that the silicon Young's Modulus may change as a function of the nanowire size and orientation[58], we assume bulk compliance coefficients hold for our nanowire devices [119]. The position vector of each atom is recalculated under given a strain condition via,

$$\mathbf{r} = (I + \varepsilon)\mathbf{r}_0, \tag{5–1}$$

where ε is the 3x3 strain tensor, $\mathbf{r}$ and $\mathbf{r}_0$ are the position vectors of the atom in the strained and unstrained crystal respectively. For a given scalar uniaxial strain along the nanowire channel direction, ϵ, the orientation specific ε can be shown as [120],

$$\varepsilon_{100} = \epsilon \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{S_{12}}{S_{11}} & 0 \\ 0 & 0 & \frac{S_{12}}{S_{11}} \end{pmatrix}, \tag{5–2}$$

$$\varepsilon_{110} = \epsilon \begin{pmatrix} \frac{1}{1+\frac{S_{44}}{2(S_{11}+S_{12})}} & \frac{\frac{S_{44}}{2(S_{11}+S_{12})}}{1+\frac{S_{44}}{2(S_{11}+S_{12})}} & 0 \\ \frac{\frac{S_{44}}{2(S_{11}+S_{12})}}{1+\frac{S_{44}}{2(S_{11}+S_{12})}} & \frac{1}{1+\frac{S_{44}}{2(S_{11}+S_{12})}} & 0 \\ 0 & 0 & \frac{2\frac{S_{12}}{(S_{11}+S_{12})}}{1+\frac{S_{44}}{2(S_{11}+S_{12})}} \end{pmatrix}$$

Table 5-1. The actual atomistic height and width of the nanowires used in this study.

	Straight ⟨110⟩	Rotated ⟨100⟩
Rounded	Atomistic	Atomistic
3nm X 3nm	3.07nm X 2.72nm	2.72nm X 2.72nm
5nm X 5nm	4.99nm X 4.75nm	4.75nm X 4.75nm
7nm X 7nm	6.91nm X 6.78nm	6.78nm X 6.78nm

where S_{ij} are the compliance coefficients. Once the atomic positions are set in the strained nanowire supercell, the device tight-binding Hamiltonian is solved for eigenvalues and eigenvectors. The strain modifies the overlap angles between neighboring atomic orbitals, which are captured by the directional cosines of the Slater and Koster two-center integrals [121]. Moreover, the strain changes interatomic bond lengths which modifies the tight-binding energy terms. The effects of bond length change on same-atom and two-center energy terms are calculated by using Boykin's modification of the generalized Harrison's rule [122, 123]. Additionally, the effect of nearest neighbor displacements due to ⟨110⟩ shear strain is captured by incorporating a generalized Lödwin's method [117]. Finally, the strained nanowire band structure is obtained for the given strain condition.

The electrostatics and self-consistency of the gate bias are introduced through Natori's compact nanowire model [112]. This model captures the effects of quantum capacitance and adjusts the nanowire band edge according to the source Fermi level, without extensive computational costs of the self-consistent solution to the 2D Poisson's relation in the device cross section and the tight-binding band structure [109, 110]. However, the gain of computational efficiency results in a loss in the accuracy in the self-consistent band structures for nanowires with larger dimensions [51, 57]. For each strain condition, the crystal band structure is calculated for each dimension and orientation, and the E-k relation is assumed to be unchanged under induced overdrive gate voltage. Using Natori's compact model [112], the inversion carrier density for a given gate overdrive voltage is found by,

$$Q = C_{eff}|V_g - V_{th}| = C_g\left[|V_g - V_{th}| - \frac{E_{0,max} - \mu_s}{q}\right] \tag{5–3}$$

where $E_{0,max}$ is the band edge energy of the ground-most valley, μ_s is the source Fermi energy, and C_{eff} is the total effective gate capacitance. The C_{eff} is modeled as the oxide (C_g) and quantum capacitances in series, without the effects of inversion and parasitic capacitances. For surrounding rectangular gates, the oxide capacitance is given by [124],

$$C_g = \frac{2\cdot\frac{5}{4}\kappa\epsilon_0}{ln\left(1+\frac{5}{4}\frac{t_{ox}}{W_{NW}}\right)} + \frac{2\cdot\frac{5}{4}\kappa\epsilon_0}{ln\left(1+\frac{5}{4}\frac{t_{ox}}{H_{NW}}\right)}, \tag{5–4}$$

where κ is the relative permittivity of the gate insulator, W_{NW} is the width, and H_{NW} is the height of the square nanowire. Simultaneously, using the semiclassical top-of-the-barrier approach [125], the carrier density of the ith band (Q_i) is found by filling the electronic states obtained from the TB band structure as a function of μ_s through,

$$Q_i = \frac{q}{\pi}\int_0^{\pi/a_x}\left\{\frac{1}{1+exp[(\mu_s - E_i(k))/k_BT]} + \frac{1}{1+exp[(\mu_d - E_i(k))/k_BT]}\right\}dk, \tag{5–5}$$

where $\mu_d = \mu_s - qV_{ds}$ is the drain Fermi level. Finally, the total carrier density is calculated by summing the contribution of all bands,

$$Q = \sum_i Q_i. \tag{5–6}$$

The self-consistency between the gate electrostatics and the band structure density-of-states (DoS) is achieved by iteratively comparing the total charge obtained from (5–3) and (5–6).

In contrast to the band edge limited Landauer's approach used in Natori's compact model for electron transport, we use a full-band approach to calculate the hole current

in the semiclassical top-of-the-barrier ballistic transport. Compared to the nanowire conduction bands, the higher complexity of the valence band structure with many local extrema is inconsistent with the assumptions used in the band edge limited current calculation [112]. The point charge is calculated for each point in the E-k relation and the point band structure group velocity is obtained via,

$$\nu(E_0, k_0) = \frac{1}{\hbar}\left(\frac{\partial E}{\partial k}\right)_{E_0,k_0}. \tag{5–7}$$

Then as shown in (5–8), the ballistic current is found through the multiplication of charge and velocity at each (E,k) point and integrated for the complete E and k ranges. This also ensures that we do not overlook the effects of the strong non-parabolicity seen in nanowire band structures. Additionally, by using the full-band current method, the potential errors in the band edge limited Landauer's approach due to the complexity of the nanowire valence bands are avoided in the ballistic hole current calculations.

$$\begin{aligned} I_{bal} = & \\ & \frac{q}{\pi\hbar}\sum_i \int_0^{\pi/a_x} \left|\frac{\partial E}{\partial k}\right| \left\{ \frac{1}{1+exp[(\mu_s - E_i(k))/k_B T]} \right. \\ & \left. - \frac{1}{1+exp[(\mu_d - E_i(k))/k_B T]} \right\} dk \end{aligned} \tag{5–8}$$

Finally, the average injection velocity is found by,

$$\langle \nu_{inj} \rangle = I_{bal}/Q \tag{5–9}$$

5.3 Results and Discussion

In this section, we present the results and discussions of the strain effects on the ballistic hole current of straight and rotated Si nanowire field-effect transistors. The ballistic hole current is calculated at a drain bias of $|V_{ds}|$=$|1V|$ and an overdrive voltage

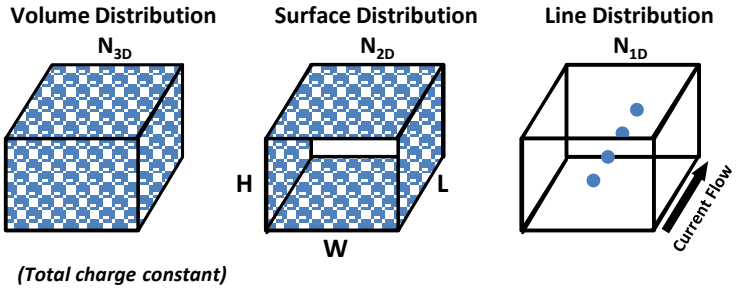


Figure 5-2. For a constant number of mobile charges in a unit volume flowing along L, different representations of the equivalent 3D, 2D and 1D charge densities are illustrated. ©2012, IEEE.

($|V_{ov}|$=$|V_{gs}$-$V_{th}|$) range of -0.4V to 1V. In order to make a fair comparison between nanowires of different sizes, the on current is normalized with the perimeter (i.e the gate width in planar devices). For benchmark purposes with conventional 2D planar devices with surface inversion, the 1D line carrier density of nanowires is converted to an equivalent 2D inversion carrier density where required, for a constant amount of mobile charge, Q, flowing along L, as shown in Figure 5-2. Without changing the total number of mobile carriers, we can convert the 1D line carrier density into 2D sidewall surface and 3D volume charge densities as shown below,

$$Q = qN_{3D}WHL = qN_{2D}(2HL + 2WL) = qN_{1D}L, \tag{5–10}$$

and thus,

$$N_{2D} = \frac{N_{1D}}{2(H + W)}, \tag{5–11}$$

where Q is the total mobile charge, q is the elementary charge, while N_{3D}, N_{2D} and N_{1D} are the equivalent 3D volume, 2D sidewall surface, and 1D line carrier densities, respectively.

The size and orientation dependent characteristics of unstrained nanowires will be discussed first. Then, the modification of these characteristics are evaluated under tensile and compressive longitudinal uniaxial strain conditions.

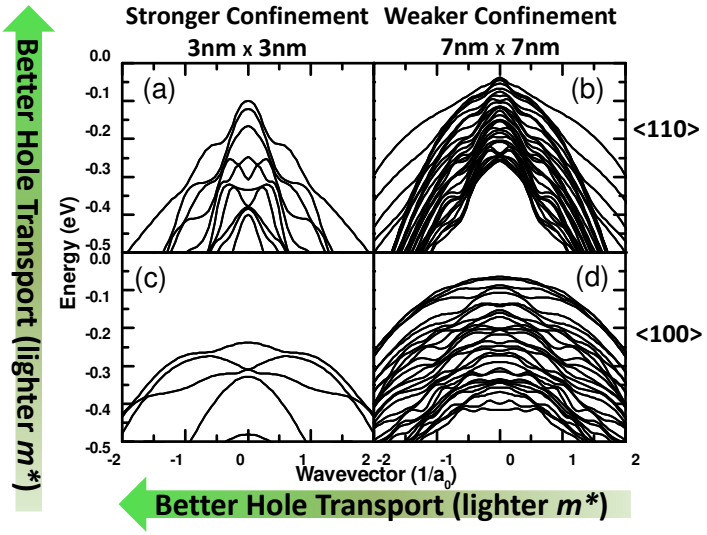


Figure 5-3. The valence band structure is shown for unstrained narrowest ($3\text{nm x } 3\text{nm}$) and widest ($7\text{nm x } 7\text{nm}$) nanowires in both $\langle 110 \rangle$ and $\langle 100 \rangle$ directions. It is seen that the hole mass decreases as we move from $\langle 100 \rangle$ direction to $\langle 110 \rangle$, as well as from a weaker structural confinement to stronger. ©2012, IEEE.

5.3.1 Unstrained Devices

The unstrained valence bandstructure of widest and narrowest nanowires are simulated for both straight and rotated devices. As shown in Figure 5-3, changes in size and orientation modify the curvature and density of the bands. Independent of crystal size, straight nanowires have larger band curvature compared to rotated devices, indicating a lighter hole effective mass [57]. As the devices scale down, the structural confinement increases. Increased quantum confinement warps and shifts the bands. As a result, narrower devices have lighter hole effective mass due to sharper bands. While lighter effective mass indicates better hole transport, it also leads to a smaller 1D DoS due to the singularity of the effective mass. In other words, unlike 2D MOSFETs, carriers in true 1D crystals have only one effective mass along the only continuous wavevector, and thus, both the transport and DoS effective mass values are the same. Consequently for narrower devices, a smaller 1D DoS is obtained due to a combination of more sparse bands and lighter effective mass.

The limited 1D DoS of nanowires coupled with thin-oxide surrounding gates result in an effective gate capacitance loss [109], similar to the "*DOS bottleneck*" shown by Fischetti *et al.* [126]. Since the effective gate capacitance is modeled as a series of oxide and quantum capacitances, an analytical expression of the quantum capacitance (C_q) can be obtained from Eq.5–3 as,

$$C_q = \frac{C_g \left[|V_g - V_{th}| - \dfrac{E_{0,max} - \mu_s}{q} \right]}{\dfrac{E_{0,max} - \mu_s}{q}}. \tag{5–12}$$

The term $(E_{0,max} - \mu_s)/q$ represents how degenerate the band structure has to become in order to support the amount of charge required by the gate electrostatics. In other words, the Fermi level has to be pushed much lower than the valence band edge due to limited DoS in the nanowire. Hence, the smaller the DoS, the larger the $(E_{0,max} - \mu_s)/q$ term is, leading to a smaller C_q. Consequently for nanowires, the C_{eff} becomes smaller than the oxide capacitance due to limited DoS, leading to a decrease in the inversion carrier density at a given overdrive voltage. For example at $V_{ov} = 0.6V$, the ratio of quantum capacitance to the oxide capacitance for the narrowest $\langle 110 \rangle$ nanowire is around 5, which leads to an almost 20% decrease in the C_{eff}. This effect becomes less pronounced for the widest $\langle 110 \rangle$ nanowire for a reduction of approximately 7%. On the other hand, the $\langle 100 \rangle$ nanowire C_q/C_g ratio is relatively insensitive to scaling at a value of around 13.5. In the limiting case, this effect is not seen for a bulk 2D inversion device for which the single gate oxide capacitance is much lower than the very large quantum capacitance due to plenty of available 2D density-of-states.

The decrease in DoS leads to a drastic enhancement in band structure carrier velocity, in contrast to the reduction of inversion carrier density. In order to compensate for the limited DoS, carriers occupy states at much higher energies in the band structure, where the Fermi velocity is high. This results in a significant increase in average carrier band structure velocity. Seen in Figure 5-4 (a), the average hole velocity for

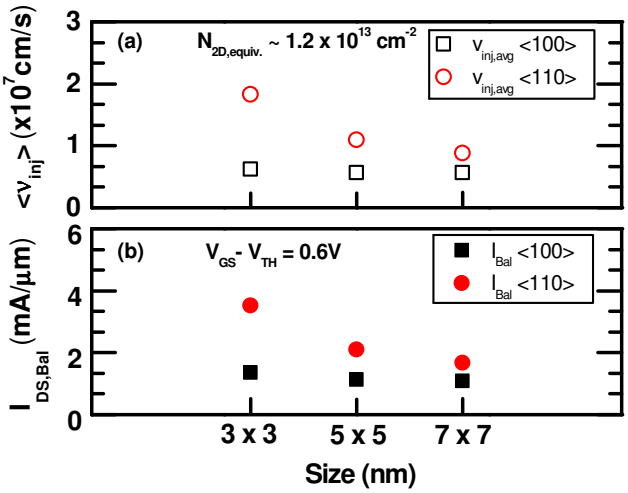


Figure 5-4. The size and orientation dependent unstrained ballistic average injection velocity (a) and hole drain current (b) are shown at an equivalent 2D carrier density of $N_{2D,eq} \sim 1.2 \cdot 10^{13} cm^{-2}$ and an overdrive voltage of $0.6V$, respectively. For a fair comparison, the drain current is normalized with transistor width (i.e. circumference). ©2012, IEEE.

$\langle 110 \rangle$ nanowires more than doubles as the size changes from the widest 7nm device ($0.89 \cdot 10^7 cm/s$) to the narrowest 3nm device ($1.83 \cdot 10^7 cm/s$). Since the negative effect of the limited DoS in total carrier density is mostly buffered with the relatively small oxide capacitance, the increase in Fermi velocity becomes the dominant factor in determining the effect of decreasing DoS for hole current. Consequently, the ballistic current, calculated as the product of velocity and carrier density, increases with decreasing DoS. The ballistic hole current is calculated at $|V_{ov}| = 0.6V$ and shown in Fig 5-4(b). In accordance with the modification of the DoS, size and orientation have direct impact on the current transport characteristics of the nanowires. While the $\langle 110 \rangle$ NWs show drastic increase in both ballistic current and velocity as the device becomes narrower, the rotated devices remain relatively insensitive to cross section scaling. These trends are in accordance with the self-consistent results of [108] and justify the choice of the compact electrostatics model method used in this work to increase the computational efficiency.

5.3.2 Strained Devices

In this subsection, the effects of strain on hole transport in ballistic nanowire devices are investigated through the modifications on nanowire band structure and DoS. Using the integrand terms in Eq. 5–5 and 5–8, we can examine the 1D band structure hole distribution ($P(E,k)$) and the 1D band structure hole current distribution ($J(E,k)$) under both strained and unstrained conditions.

The strain-induced modifications of the band structure and DoS are simulated for the strained 3nm X 3nm $\langle 100 \rangle$ nanowire device. The band structure and $P(E,k)$ are shown for different strain conditions in Figure 5-5, at an overdrive voltage of $0.6V$. Only the positive wavevector portion of the complete band structure is plotted since the E-k dispersion relation is an even function of wavevector k. As seen from the figure, uniaxial compression increases the energy separation between the top and the second valence bands, while warping the top band into a sharper dispersion relation. On the other hand, tension lowers the band splitting energy and shapes the top band into a flatter dispersion. Consequently, the 1D DoS around the valence band maximum decreases with compression and increases with tension. This leads to a lower inversion carrier density for a compressively strained device, with holes mostly localizing near the light mass region at the top valence band. Contrarily, for tensile stressed devices, the increase in DoS leads to higher inversion carrier density which are now mostly residing at a flatter band structure dispersion. Hence, uniaxial stress leads to opposite sign modifications of total hole carrier density and band structure velocity in 1D nanowires. This effect is not observed in conventional MOSFETs, for which the changes in inversion carrier density due to the uniaxial strain-induced modifications in DoS is shown to be negligible for bulk 2D inversion devices [127].

Using the product of the hole distribution and the $\nu(E,k)$ at a given point in the strained band structure, we can investigate the strain modification of the hole current. The effect of strain on 3nm X 3nm $\langle 100 \rangle$ nanowire $J(E,k)$ can be seen in Figure 5-6.

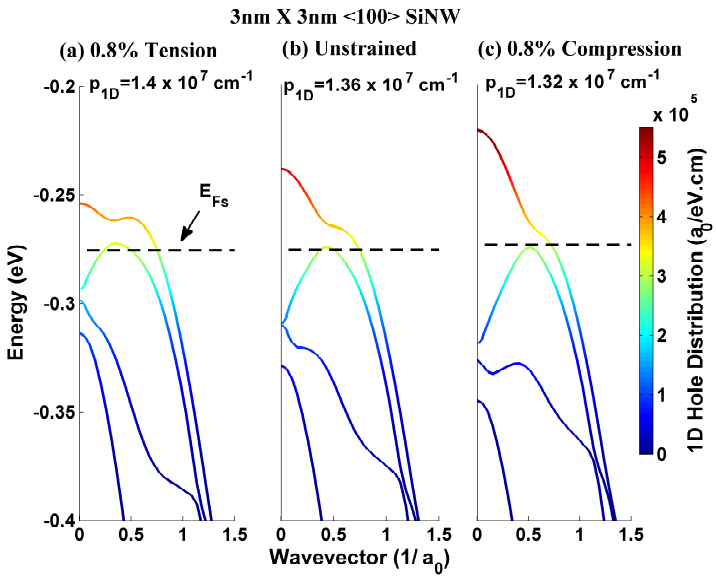


Figure 5-5. (Color online) The band structure hole distribution is shown for 3nm X 3nm ⟨100⟩ nanowire under (a) 0.8% tension, (b) unstrained and (c) 0.8% compression conditions at $|V_{ov}| = 0.6V$. ©2012, IEEE.

Compared to the unstrained case, under 0.8% compression, the increase in the portion of total carriers residing at the steep slope region of the top valence band results in more than 8% increase in current, despite the reduction of inversion carrier density by an amount of 3%. On the other hand for tensile strain, an increase of 3% in hole density is not enough to compensate for the loss in group velocity at the much flatter top region of the valence band, leading to an 8% loss in ballistic hole current. Thus, the effect of strain-induced reduction in DoS for ballistic injection velocity improvement in true 1D nanowires is in accordance with the expectations stated in Ref. [106].

Similar to the ⟨100⟩ devices, compressive stress is also beneficial for hole transport in ⟨110⟩ nanowires. As we have mentioned before, even in the unstrained case, the ⟨110⟩ nanowire band structure is much more desirable for hole transport than nanowires patterned in the ⟨100⟩ direction. Additionally, ⟨110⟩ nanowires are more sensitive to uniaxial strain compared to ⟨100⟩ devices. This is explained by the greater deformation of the crystal symmetry by the shear component of the ⟨110⟩ strain tensor, leading to significant changes in the band structure [2]. The strain effects on 3nm X 3nm ⟨110⟩

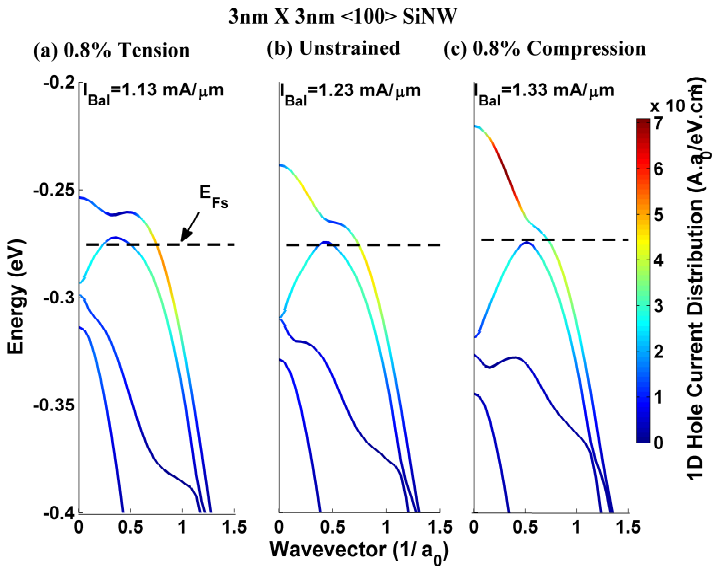


Figure 5-6. (Color online) The band structure current distribution is shown for 3nm X 3nm $\langle 100 \rangle$ nanowire under (a) 0.8% tension, (b) unstrained and (c) 0.8% compression conditions at $|V_{ov}| = 0.6V$. ©2012, IEEE.

nanowire $P(E, k)$ and $J(E, k)$ are shown in Figure 5-7 and Figure 5-8. Under 0.8% uniaxial compression, almost all of the holes are repopulated to the topmost valence band with decreased effective mass due to band warping. The result is a 14% increase in ballistic current. For devices with 0.8% tensile strain, holes occupy the heavier mass regions of the band structure, leading to a total current decrease of almost 20%.

In order to normalize the device performance enhancement for NWs with different size and orientations, the gauge factor (GF) is calculated for all devices using Eq. 5–13. The GF, normalized change in resistance per unit strain, for a semiconductor is obtained by,

$$GF = \frac{R^s - R^0}{R^0}\frac{1}{\epsilon} = \frac{I^0 - I^s}{I^s}\frac{1}{\epsilon}, \tag{5–13}$$

where R^s and R^0 are the strained and unstrained channel resistances, while I^s and I^0 are the strained and unstrained drain currents respectively. The gauge factors for nanowires are calculated at 0.8% compression and $V_{ov} = 0.6V$, and shown in Figure 5-9. The $\langle 110 \rangle$ NW GF is found to be much larger compared to the $\langle 100 \rangle$ devices, because the shear strain component of the $\langle 110 \rangle$ uniaxial stress effectively breaks the

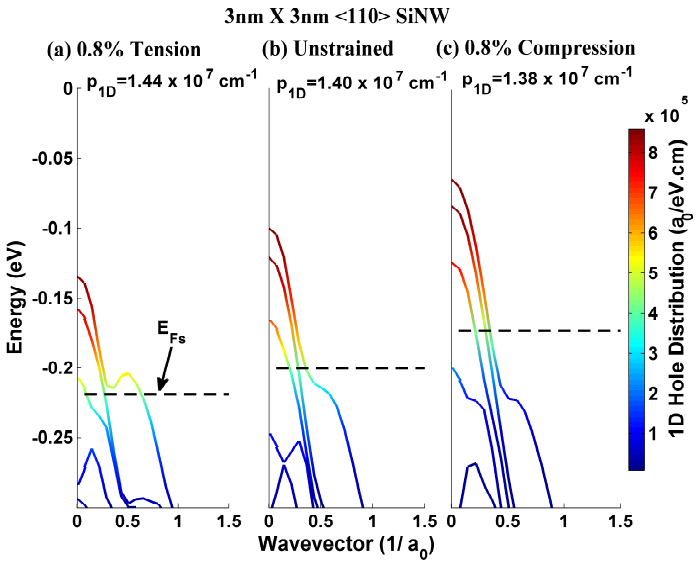


Figure 5-7. (Color online) The band structure hole distribution is shown for 3nm X 3nm $\langle 110 \rangle$ nanowire under (a) 0.8% tension, (b) unstrained and (c) 0.8% compression conditions at $|V_{ov}| = 0.6V$. ©2012, IEEE.

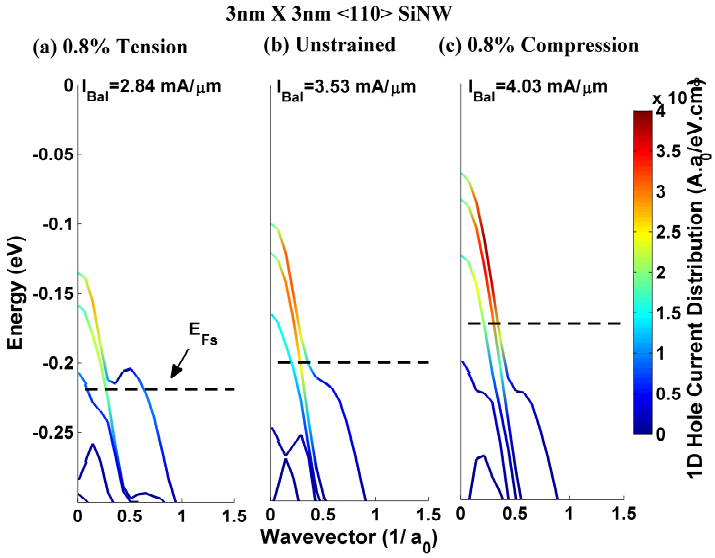


Figure 5-8. (Color online) The band structure current distribution is shown for 3nm X 3nm $\langle 110 \rangle$ nanowire under (a) 0.8% tension, (b) unstrained and (c) 0.8% compression conditions at $|V_{ov}| = 0.6V$. ©2012, IEEE.

crystal symmetry. It is also seen that the GF for $\langle 110 \rangle$ nanowires decreases as the nanowire cross section dimensions shrink. This is explained by a mechanism similar to the smaller strain-induced enhancement of 2D MOSFET electron low-field mobility at high inversion carrier densities. Both increased spatial confinement and uniaxial

compression enhances the ⟨110⟩ NW hole transport via a reduction in the DoS. Hence, the strain-induced current improvement for the narrowest 3nm ⟨110⟩ nanowire is less than wider 5nm and 7nm devices. On the other hand, the widest 7nm ⟨100⟩ device has a comparable trade-off between strain-induced inversion carrier density loss and group velocity increase, and hence has a very low GF. As the ⟨100⟩ devices shrink, the velocity enhancement increases while the carrier density loss remains almost constant, leading to higher GF for downscaled ⟨100⟩ nanowires.

In addition to the drive current enhancement due to strain effects on the band structure, mechanical strain may further improve transistor performance if employed during nanowire fabrication to reduce surface atomic disorders. It is observed through molecular dynamics studies that very high levels of radial compression is memorized in the nanowire core, as a result of the thermal oxidation processes to thin the nanowire channel [128]. Recently, it is shown that this radial compression leads to a large scale of spatial disorder at the near surface atomic locations in nanowires subjected to thermal oxidation [129]. This in return results in localized states within the bands and increased carrier mass, especially for narrow nanowires with high surface to volume ratio. Therefore, a significant amount of (up to 2 times) reduction in ballistic current is expected in nanowire transistors due to atomic disorder [129]. In this case, a wafer level uniaxial 4-point bending technique can be used to induce uniaxial compression along the nanowire channel during the thermal oxidation process to relax the radial compression build up. Consequently, the atomic disorder at the interface and the related current degradation may be minimized.

5.4 Summary

The underlying physics of strain effects on ballistic hole transport in nanowire field-effect transistors are explained by the strain-induced modification of the density-of-the-states for the first time. The strain response of ballistic p-type NWs is investigated as a function of channel orientation and cross section dimensions. It is found that the

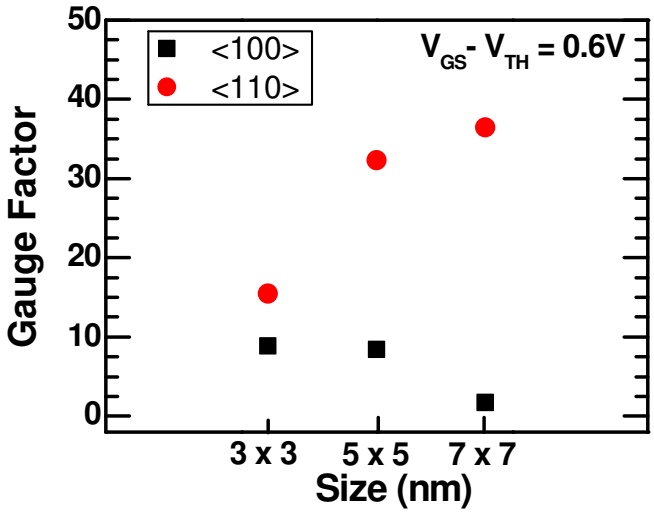


Figure 5-9. Size dependent gauge factor is shown for both straight and rotated p-type nanowires at an overdrive voltage of 0.6V. ©2012, IEEE.

⟨110⟩ nanowires respond to uniaxial strain more than their ⟨100⟩ counterparts, as the shear strain component in ⟨110⟩ stress breaks the crystal symmetry and leads to significant changes in band structure. Also for ⟨110⟩ nanowires, the effects of compressive strain and structural confinement are shown to be similar, which results in a decrease in the strain-induced enhancement of ballistic hole current at smaller dimensions. It is also shown that unlike the bulk 2D inversion devices, the reduction of the nanowire 1D DoS under uniaxial compression decreases the total inversion hole density and has a negative effect on the total current enhancement. Nevertheless, the high unstrained ballistic drive current capability and large improvement with uniaxial compression indicate that narrow ⟨110⟩ nanowire p-type field-effect transistors are potential candidates for high-performance logic devices, while wider ⟨110⟩ devices with larger GF are expected to be suitable for sensor applications. Finally, it is suggested that introducing a wafer level uniaxial tension during the thermal oxidation process may improve the drive current of nanowire transistors by reducing the atomic disorders at the interface.

CHAPTER 6
PHYSICAL INSIGHTS ON STRAIN ENHANCED SURFACE ROUGHNESS LIMITED ELECTRON MOBILITY IN SILICON INVERSION LAYERS

6.1 Motivation

The deviation of the silicon/oxide boundary from an ideal interface and its effects on carrier transport in silicon inversion layers have been thoroughly investigated since the early 1980s' pioneering works by Goodnick *et al.* [28] and Ando *et al.* [130]. The statistical distribution of the silicon/oxide interface position introduces a perturbation in the self-consistent potential, which results in the scattering of MOSFET inversion carriers near the surface. Especially for contemporary planar devices with very high levels of doping ($\sim 10^{17} - 10^{18} cm^{-3}$) and operating at high inversion carrier densities ($\sim 10^{13} cm^{-2}$), the vertical electric field which confines the carriers in the quantum well near the surface goes up to the order of $\sim 1 MV/cm$. At these very high field conditions, carriers reside significantly close to the oxide, resulting in very high surface roughness scattering rates. In fact, even at room temperature, surface roughness scattering is the dominant scattering mechanism that limits mobility at "ON" state [131, 132]. Therefore, understanding the effects of mechanical stress on the surface roughness limited carrier mobility is crucial to model and predict the performance of modern day strained logic devices.

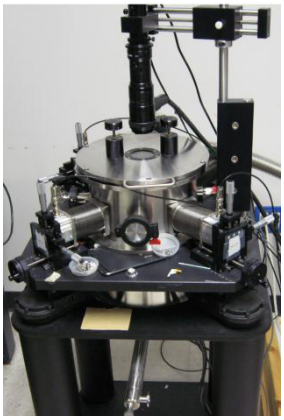
One way to investigate the strain effects on surface roughness limited mobility is to measure the low-field effective mobility at low temperatures under different strain conditions. Since at low temperatures phonon scattering becomes less pronounced and carriers are further confined to the silicon/oxide interface, surface roughness scattering becomes the dominant limiting mechanism for low-field electron mobility. Hence, the resultant strain-induced change in linear drain current or mobility can be mostly attributed to the modification in surface roughness limited mobility. The effects of stress on surface roughness limited mobility have been systematically investigated for p-type single-gate MOSFETs [133, 134]. It was found that the increased strain-induced

enhancement of hole mobility at lower temperatures is a result of the greater impact of band warping on hole effective mass reduction. However, it is known that electrostatic confinement induced valley splitting in silicon conduction band leads to almost complete repopulation of electrons to lower Δ_2 valley with lighter effective mass. Thus at high vertical fields, further strain-induced reduction of electron mass cannot be the reason of electron mobility enhancement alone.

In order to understand the discrepancy between the theoretical expectations and experimental results of surface roughness mobility in strained nMOSFETs, an *ad hoc* strain-induced modification of surface morphology has been proposed by Fischetti *et al* [29]. It is found under biaxial tensile stress that by decreasing the rms value of the roughness amplitude and increasing the correlation length of the roughness distribution, the model can accurately predict the experimental results of enhanced electron mobility at high inversion carrier density [29]. Later, Bonno *et al.* provided further experimental support to the *ad hoc* modification of the surface roughness via AFM measurements on epitaxially 0.8% strained SOI devices [135]. Additional experimental support to Fischetti's model has been published by Zhao *et al.* through the characterization of the effects of process induced biaxial strain on surface roughness mobility at cryogenic temperatures [136, 137].

To extend the systematical approach employed for pMOSFETs, we have measured the strain-induced linear drain current enhancement as a function of temperature, while applying controlled longitudinal uniaxial stress using a 4-point bending setup. Also, to characterize the morphological changes with strain, a similar 4-point bending setup is used to control the magnitude of applied stress during AFM measurements.

The underlying physics of increased uniaxial strain-induced enhancement of linear drain current in nMOSFETs are investigated by a self-consistent MOSFET simulator and a mobility model using Kubo-Greenwood formula with varying surface roughness conditions under stress.



Lakeshore Cryogenic Station
(a)

Low Temperature Stress Jig
(b)

Figure 6-1. (a) The cryogenic temperature wafer bending bending setup is shown on the left and (b) the stress jig used in the cryogenic station is shown on right.

6.2 Experimental

In order to investigate the stress effects on surface roughness limited mobility in n-type silicon inversion layers, the linear drain current is measured under uniaxial tensile stress conditions at low temperatures, for which the influence of phonon scattering is reduced. The low temperature strain measurements were conducted on a custom built Lakeshore systems cryogenic station with liquid nitrogen (77K) and helium capabilities (4K). The cryogenic station and the low temperature stress jig are shown in Figure 6-1.

Long channel ($L_{gate} = 10\mu m$) nMOSFET linear drain current is measured at $V_{DS} = 50mV$ for a temperature range varying from 293K down to 80K. At these temperature conditions, uniaxial tensile stress is applied and the strain-induced linear drain current enhancement is measured at an overdrive voltage of $0.5V$. The linear drain current enhancement as a function of temperature is shown in Figure 6-2. At room temperature, the strain-induced linear drain current enhancement is measured to be equivalent to the piezoresistance coefficient of (100)/$\langle 110 \rangle$ nMOSFETs [104]. As the temperature decreases down close to the liquid nitrogen level, the strain enhancement of linear drain current monotonically increases.

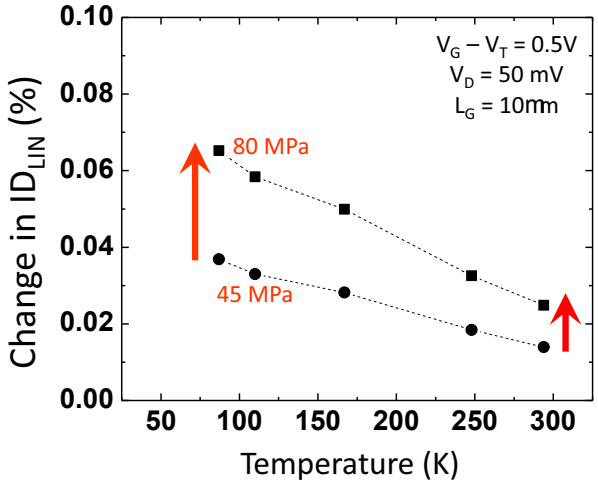


Figure 6-2. The strain enhancement of nMOSFET linear drain current is shown as a function of sample temperature for two different stress conditions.

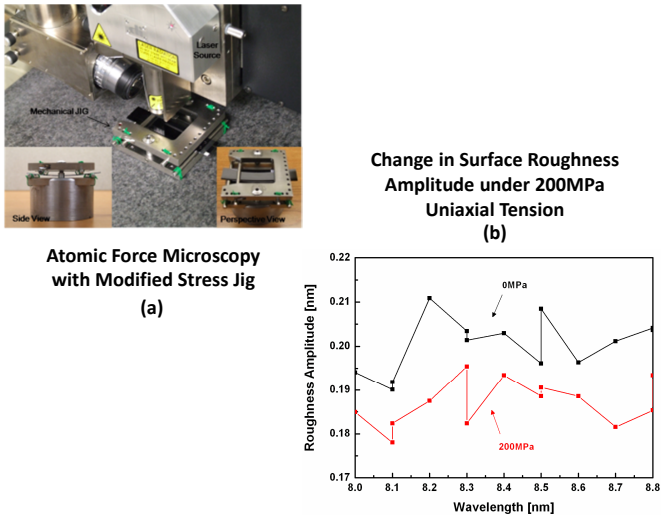


Figure 6-3. (a) AFM system is shown with the mechanical bending jig. (b) The modification of the roughness amplitude is shown at 200MPa uniaxial tensile stress. [138]

At the same time, atomic force microscopy (AFM) scan of the bare silicon surface is done under mechanical stress just after the removal of the native oxide [138]. As shown in Figure 6-3, it is found that uniaxial tensile stress results in a decrease in the roughness amplitude of the bare silicon surface.

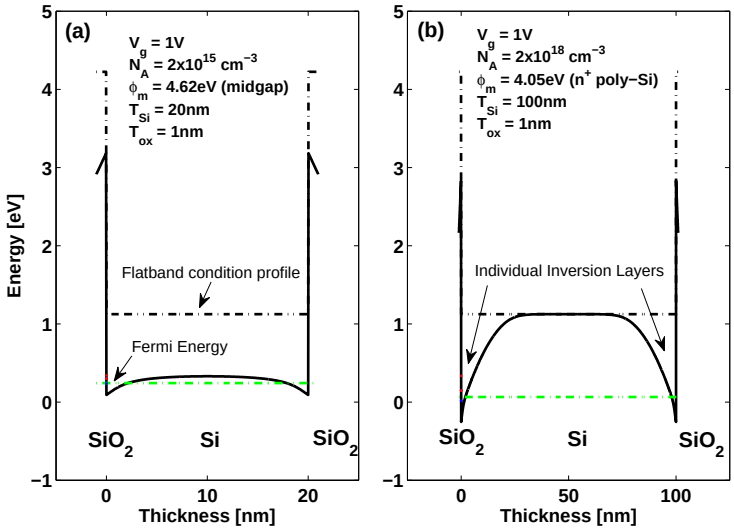


Figure 6-4. (a) Double-gate operation of undoped 20nm thick (100) MOSFET conduction band profile is shown. (b) With high doping level and 100nm thick silicon, each gate is isolated and operates as individual single-gate MOSFETs.

6.3 Simulation

In order to simulate the strain effects on surface roughness limited mobility in silicon inversion layers, we employ a modified version of the self-consistent double gate FinFET simulator developed in Chapter 4. By setting the separation between two gates larger than the cumulative depletion width of both gates, we can effectively isolate the individual inversion layers under each gate and obtain a single gate operation for two parallel distinct MOSFET structures. The conduction band profiles of the metal/oxide/silicon heterostructure for single-gate MOSFET and double-gate FinFET operations are shown in Figure 6-4. For single-gate operation, the potential deep inside the fin goes to zero, indicating a clear separation of each inversion layer. On the other hand for thin fins with low doping density, the inversion layers merge at the center of the fin and result in a non-zero potential profile along the fin thickness.

6.3.1 Temperature Dependence

The temperature effects are captured within the self-consistent Schrödinger-Poisson simulator through the value of the thermal voltage $k_B T/q$ and the temperature

dependent silicon band gap energy, bulk conduction, and valence band effective density-of-states. The temperature dependent silicon band gap is calculated by [139],

$$E_g(T) = 1.17 - 4.73 \cdot 10^{-4} \frac{T^2}{T + 636}. \tag{6–1}$$

Similarly, the conduction and valence band effective density of states as a function of lattice temperature are given as,

$$\begin{aligned} N_{c,eff}(T) &= 6.2 \cdot 10^{15} T^{3/2} [cm^{-3}] \\ N_{v,eff}(T) &= 3.5 \cdot 10^{15} T^{3/2} [cm^{-3}]. \end{aligned} \tag{6–2}$$

Using Eq. 6–1 and 6–2, we can calculate the temperature dependent intrinsic carrier density in silicon,

$$n_i(T) = \sqrt{N_{c,eff}(T) N_{v,eff}(T)} exp(\frac{-E_g(T)}{2 k_B T}) [cm^{-3}]. \tag{6–3}$$

6.3.2 Strain Modification

Energy band splitting and band warping effects of mechanical stress on the silicon conduction band is calculated using a deformation potential based model[140]. In this model, a parabolic conduction band valley *E-k* dispersion relation is given by,

$$\begin{aligned} E_\nu = \frac{\hbar^2}{2}(\frac{k_x^2}{m_x^\nu} + \frac{k_y^2}{m_y^\nu} + \frac{k_z^2}{m_z^\nu}) + \Xi_d(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}) \\ + \Xi_u \epsilon_{ii} + \Xi_m \epsilon_{jk} k_j k_k, \end{aligned} \tag{6–4}$$

where ν denotes the doubly degenerate X-Y-Z valleys, Ξs are the deformation potentials [141], ϵ is the mechanical strain, subscript i is the direction along the major axis of the valley ν, and finally, subscripts j,k denote the out-of-plane direction for the same valley [142]. Using a 3x3 strain tensor as in Eq. 5–2, we can calculate the strain-induced

energy shifts in each of the X, Y and Z valleys by,

$$\begin{aligned} \Delta E_X &= \Xi_d(\epsilon_{XX}+\epsilon_{YY}+\epsilon_{ZZ})+\Xi_u\epsilon_{XX}, \\ \Delta E_Y &= \Xi_d(\epsilon_{XX}+\epsilon_{YY}+\epsilon_{ZZ})+\Xi_u\epsilon_{YY}, \\ \Delta E_Z &= \Xi_d(\epsilon_{XX}+\epsilon_{YY}+\epsilon_{ZZ})+\Xi_u\epsilon_{ZZ}. \end{aligned} \tag{6–5}$$

During each pass of the self-consistent loop, the strain-induced energy shifts of conduction band valleys are superposed on the subband energies obtained from the Schrödinger's equation. The occupancy of each subband ladder in each valley is calculated with the superposed valley energies and the resulting spatial carrier distribution is input in the Poisson solver. With the incorporation of the band shift effects of mechanical strain, the strained eigenenergies and eigenvectors are self-consistently calculated.

Similarly, the effect of shear strain induced band warping is obtained via the $\Xi_m\epsilon_{jk}k_jk_k$ term in Eq. 6–4. For (100)/$\langle$110$\rangle$ inversion layers, the tensile shear strain decreases the transport mass of Δ_2 valley. The band warping dependent transport effective mass of Δ_2 is given by [143],

$$m^*_{\Delta_2,110} = \frac{m_t}{1+\alpha\epsilon_{XY}m_t}, \tag{6–6}$$

where α is a constant experimentally found to be $86.8/m_0$ [144].

6.3.3 Mobility Model

Once the self-consistent valley energies and wavefunctions are found for the given bias, temperature and strain conditions, the low-field effective electron mobility is calculated using the Kubo-Greenwood formula [145, 146]. Since we are interested in the strain effects at high inversion carrier densities, the Coulomb scattering potential of ionized impurities and interface traps is assumed to be effectively screened by mobile electrons. Hence, the mechanisms chosen to contribute to electron scattering are acoustic phonon scattering (μ_{ac}), optical phonon scattering (μ_{op}), and surface roughness

scattering (μ_{sr}). Based on the energy and wavevector transfer characteristics between the scattering potential and the electron, the transition of the carrier may happen within the same valley (intravalley) or between different valleys (intervalley). The cumulative scattering rate of different mechanisms is found by the Mathiessen's Rule,

$$\frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3} + \dots , \tag{6–7}$$

where μ_n's are mobility values resulting from and limited by independent scattering events.

6.3.3.1 Intravalley phonon scattering

Electrons traveling through the semiconductor channel interact with quantized lattice vibrations (phonons). As a result of this interaction, electrons may scatter to a different state within the same valley via an elastic intravalley scattering event. For cubic semiconductors, elastic intravalley transitions via optical phonons are neglected and only acoustic phonons are considered [147]. The elastic nature of the acoustic phonon interaction results in changes in electron wavevector without altering the scattered electron's energy. Due to the short wavevectors of the acoustic branch of the phonon dispersion, acoustic phonon interaction of electrons only result in scattering within the same valley. However, scattered electrons may transfer within the subbands of the initial valley. Thus, acoustic phonon scattering is an intravalley/intersubband process. The influence of the acoustic phonons on electron transport in silicon inversion layers is modeled using an isotropic deformation potential approximation for the elastic intravalley scattering process. The phonon number is found by the equipartition approximation. For an electron at the ith subband of valley ν, the energy dependent acoustic scattering rate is the summation of all possible scattering events to the jth subband of the same valley. The intravalley acoustic phonon scattering rate is given by [148],

$$\frac{1}{\tau_{ac}^{\nu,i}(E)} = \sum_j \frac{m_d D_{ac}^2 k_B T}{\hbar^3 \rho s_l^2} F_{i,j}^{\nu} U(E - E_j^{\nu}), \tag{6–8}$$

where m_d is the density of states effective mass for the valley ν, D_{ac} is the acoustic deformation potential, and ρ and s_l are the mass density and longitudinal speed of sound of the crystal, respectively. The term $F_{i,j}^{\nu}$ denotes the "electron-phonon wavefunction overlap" or the "form factor" of the scattering event and is given by,

$$F_{i,j}^{\nu} = \int \psi_{i,\nu}^2(z)\psi_{j,\nu}^2(z)\,dz, \tag{6–9}$$

where $\psi_{i,\nu}$ and $\psi_{j,\nu}$ are the envelope functions of the ith and j^{th} subbands along the confinement axis, $U(E)$ is the Heaviside step function, and E_j is the energy of the subband.

6.3.3.2 Intervalley phonon scattering

In addition to intravalley transitions by acoustic phonons, electrons may also scatter to different valleys. The intervalley transitions are assisted by large wavevector acoustic and optical phonons at relatively high energies. This leads to inelastic scattering of electrons. During an inelastic scattering event, electrons may lose energy while the lattice emits a phonon, or electrons may absorb energy from the suppression of a phonon within the lattice. Depending on the angle of the electron wavevector change, intervalley scattering events can be *f-type* or *g-type* transitions [147, 149]. Transverse optical, transverse acoustic and longitudinal acoustic phonons may assist an *f-type* transition. Similarly, for a *g-type* scattering event, longitudinal optical, transverse acoustic and longitudinal acoustic phonons are involved. For an electron at the ith subband of valley ν, the energy dependent intervalley phonon scattering rate is the summation of all possible scattering events to the jth subband of the valley μ with the assistance of the phonon m. The intervalley phonon scattering rate is given by [150],

$$\frac{1}{\tau_{iv}^{\nu,i}(E)} = \sum_{\mu}\sum_{j}\sum_{m} \frac{n^{\nu,\mu} m_d^{\mu,j} D_{iv,m}^2}{2\hbar\rho E_{iv,m}} F_{i,j}^{\nu,\mu} \quad (N_{iv} + \tfrac{1}{2} \pm \tfrac{1}{2}) \frac{1 - f(E \mp E_{iv,m})}{1 - f(E)} \\ U(E \mp E_{iv,m} - E_j^{\mu}), \tag{6–10}$$

where $n^{\nu,\mu}$ is the degeneracy level of the process, $m_d^{\mu,j}$ is the density of states effective mass for the final valley, $D_{iv,m}$ is the intervalley deformation potential for phonon m, $E_{iv,m}$ is the energy of the intervalley phonon m, $N_{iv,m}$ is the Bose-Einstein occupation number of the intervalley phonon mode of energy $E_{iv,m}$, $(N_{iv,m} + \frac{1}{2} \pm \frac{1}{2})$ is the modification for phonon emission and absorption, and $U(E)$ is the Heaviside step function for a given energy. The selection rules for allowed intervalley transitions can be found in [149].

6.3.3.3 Surface roughness scattering

As the electrons in silicon inversion layers are confined to a very narrow quantum well at the silicon/SiO_2 interface, the irregularities at the crystal-oxide boundary creates a perturbation to the self-consistent potential, which leads to surface roughness scattering (SRS). The surface roughness scattering is an elastic process which results only in the modification of the electron wavevector. Hence, surface roughness scattering is an intravalley/intersubband process.

For a deviation from the ideal interface boundary ($\Delta(\mathbf{r})$) at the coordinate $\mathbf{r}$, the potential near the interface is given by [151, 152],

$$V(z + \Delta(\mathbf{r})) = V(z) + \Delta(\mathbf{r})\frac{\partial V(z)}{\partial z}, \tag{6–11}$$

where $V(z)$ is the unperturbed potential and z is the axis normal to the gate plane. Using Eq. 6–11, the perturbation potential at location (z,$\mathbf{r}$) can be approximated by,

$$H_{SR}(z, \mathbf{r}) = e\Delta(\mathbf{r})E(z), \tag{6–12}$$

where $E(z)$ is the transverse electric field. Then, the scattering matrix element from $\imath^{th}$ subband to $\jmath^{th}$ subband is given by,

$$|M_{i,j}(\mathbf{q})| = e^2 \left| \int \psi_i(z)E(z)\psi_j(z)dz \right|^2 |\Delta(\mathbf{q})|^2, \tag{6–13}$$

where $\mathbf{q} = \mathbf{k} - \mathbf{k'}$ is the difference between the initial and scattered wavevectors, and $\Delta(\mathbf{q})$ is the Fourier transform of the $\Delta(\mathbf{r})$, which is described by statistical formulae

rather than deterministic models [28, 130]. As seen in Eq. 6–13, the Fourier transform of the autocovariance function of the deviation from the ideal interface is employed for carrier transport analysis in the wavevector space. Within the literature, it is commonly accepted that the deviation from the ideal interface can be approximated by either a Gaussian or an exponential distribution function whose Fourier transform is given by,

$$|\Delta(\mathbf{q})|^2 = \frac{\pi \Delta_m^2 L_m^2}{(1 + q^2 L_m^2/2)^{1+n}}, \tag{6–14}$$

where Δ_m is the *rms* amplitude of the surface roughness and L_m is the correlation length of the distribution. The parameter n determines the characteristics of the distribution and is generally taken as $n = 2$ [153] or $n = 1/2$ [152], where the latter value results in a well agreed exponential distribution suggested by Ando, Fowler and Stern [28]. Using the perturbation Hamiltonian (H_{SR}) and the well accepted exponential roughness distribution, the surface roughness scattering rate for an electron in the ith subband to the jth subband is given by [151],

$$\frac{1}{\tau_{SR}^{\nu,i}(\mathbf{k}_{\|})} = \frac{m_d^{\nu,j} e^2 E_{eff}^2 \Delta_m^2 L_m^2}{2\hbar^3} \int_0^{2\pi} \frac{d\theta}{\epsilon(q)\left(1 + \frac{L_m^2 q^2}{2}\right)^{3/2}}, \tag{6–15}$$

where $\epsilon(q)$ is the static dielectric function with screening effects from mobile carriers included. The effects of screening is considered only for intrasubband transitions and the screened static dielectric function is given by,

$$\epsilon(q) = 1 + \frac{e^2}{2\kappa_{Si}\epsilon_0} \frac{1}{q} \frac{m_d^{\nu,j}}{\pi\hbar^2} F(q), \tag{6–16}$$

for which $F(q)$ is the screening overlap factor and is calculated through,

$$F(q) = \sum_i \int dz \int dz' |\psi_i(z)|^2 |\psi_i(z')|^2 e^{(-q|z-z'|)}. \tag{6–17}$$

While for bulk single-gate MOSFETs it is shown in Eq. 6–12 that the surface roughness perturbation potential can be estimated by extending the ideal boundary

electric field along the displacement of the rough interface, this approach leads to an inaccurate representation of the H_{SR} for ultra-thin-body single-gate MOSFETs [152] and double-gate FinFETs [154]. In order to calculate the surface roughness perturbation potential accurately in non-bulk devices, the approximation in Eq. 6–12 is replaced by the generalized definition of the potential perturbation given by [152],

$$H_{SR}(z,\mathbf{r}) = -e\frac{\Delta(\mathbf{r})\Delta V_m(z)}{\Delta_m}, \tag{6–18}$$

where $\Delta V_m(z)$ is the difference between the self-consistent potentials obtained for the ideal interface condition and for the condition of ideal interface displaced by Δ_m. Consequently, the surface roughness scattering rate for non-bulk planar devices is finally given by,

$$\frac{1}{\tau_{SR}^{\nu,i}(\mathbf{k}_{\|})} = \frac{m_d^{\nu,j} e^2 \Delta_m^2 L_m^2}{2\hbar^3} \left| \int \psi_i(z)\frac{\Delta V_m(z)}{\Delta_m}\psi_j(z)\,dz \right|^2 X \int_0^{2\pi} \frac{d\theta}{\epsilon(q)\left(1+\frac{L_m^2 q^2}{2}\right)^{3/2}}. \tag{6–19}$$

6.4 Results and Discussion

The room temperature (300K) low-field electron mobility is calculated for an unstrained 10μm long single-gate, 10μm wide (100)/$\langle 110\rangle$ nMOSFET device with $N_A = 2\cdot 10^{18} cm^{-3}$ doping density. During the mobility calculation, the acoustic phonon deformation potential is taken as $12eV$ [155] and the intervalley phonon energies, deformation potentials, and selection rules are taken from [149]. The surface roughness parameters are taken as $\Delta_m = 0.4nm$ and $L_m = 2.4nm$ from [133], where it is shown that these values accurately fit the pMOSFETs taken from the same CMOS wafer lot with the nMOSFETs used in this study. The summary of the mobility parameters can be seen in Table 6-1. The simulated low-field effective mobility and its constituents are shown with respect to effective electric field (E_{eff}) in Figure 6-5. As the field increases,

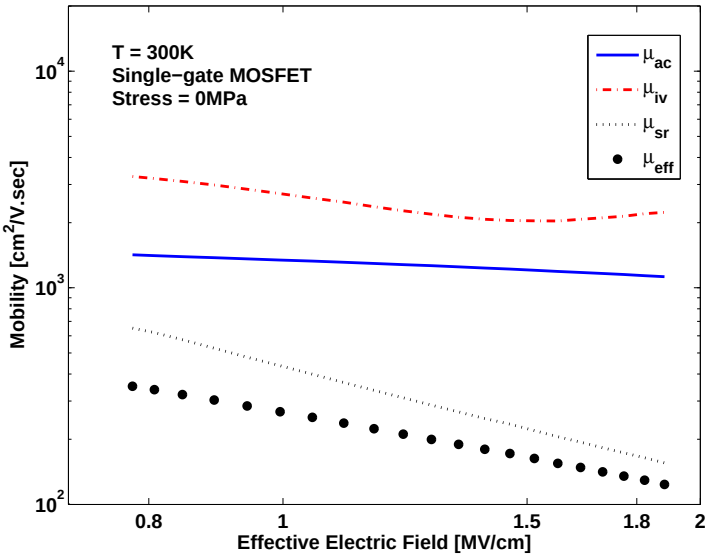


Figure 6-5. The effective mobility and its constituents are plotted vs. effective electric field at 300K and no stress. Almost complete log-log dependence of total mobility to field indicates the significance of roughness scattering even at room temperature.

the intervalley phonon limited mobility decreases with the increase in the occupancy of the lowest subband of the Δ_4 valley, which has a higher intervalley scattering rate than the ground subband of Δ_2. However, a further increase in the field leads to larger splitting of Δ_4^0 and Δ_2^1, and the overall intervalley phonon scattering decreases. At the same time, a higher field results in further confinement of carriers to a narrower quantum well. This results in an increased form factor and the intravalley acoustic phonon limited mobility monotonically decreases at higher fields. Similarly at higher electric fields, the stronger confinement of electrons towards the interface drastically decreases the surface roughness scattering limited mobility. In fact, the total effective low-field mobility calculated for a gate bias range of $0.3V$ to $1.3V$ shows almost linear log-log dependence with the field, which indicates the significance of surface roughness scattering even at room temperature.

The room temperature electron mobility of a 20nm thick undoped ($N_{A,DG} = 2 \cdot 10^{15} cm^{-3}$) double-gate MOSFET with midgap-metal ($\phi_m = 4.62eV$) is also simulated

Table 6-1. The values of phonon deformation potentials [149, 155], phonon energies [149] and surface roughness parameters [133] used for mobility calculations are shown.

Parameter	Symbol	Value	Units
Acoustic phonon deformation potential	D_{ac}	12	eV
Silicon crystal density	ρ	2.329	g/cm^3
Silicon sound velocity	s_l	$9.037 \cdot 10^4$	cm/s
f-type TA phonon deformation potential	$D_{iv}^{(f,TA)}$	0.3	$10^8 eV/cm$
f-type LA phonon deformation potential	$D_{iv}^{(f,LA)}$	2	$10^8 eV/cm$
f-type TO phonon deformation potential	$D_{iv}^{(f,TO)}$	2	$10^8 eV/cm$
g-type TA phonon deformation potential	$D_{iv}^{(g,TA)}$	0.5	$10^8 eV/cm$
g-type LA phonon deformation potential	$D_{iv}^{(g,LA)}$	0.8	$10^8 eV/cm$
g-type LO phonon deformation potential	$D_{iv}^{(g,LO)}$	11	$10^8 eV/cm$
f-type TA phonon energy	$E_{iv}^{(f,TA)}$	19	meV
f-type LA phonon energy	$E_{iv}^{(f,LA)}$	47.4	meV
f-type TO phonon energy	$E_{iv}^{(f,TO)}$	59	meV
g-type TA phonon energy	$E_{iv}^{(g,TA)}$	12	meV
g-type LA phonon energy	$E_{iv}^{(g,LA)}$	18.5	meV
g-type LO phonon energy	$E_{iv}^{(g,LO)}$	61.2	meV
Surface roughness amplitude	Δ_m	0.4	nm
Surface roughness correlation length	L_m	2.4	nm

for gate bias conditions of up to 1.3V. The comparison of electron mobility with respect to the single-gate bulk MOSFET is seen in Figure 6-6 and 6-7. For the same inversion electron density, the undoped double-gate mobility is twice the mobility for the doped bulk single-gate device (Figure 6-6). As seen in the Figure 6-7, the effective electric field is much lower for the undoped double-gate device, which greatly increases the surface roughness limited mobility. In fact, up to $V_G = 0.8V$, the main limiting mechanism for the double-gate transistor is the phonon scattering. However, at high field conditions (close to the doped single-gate operating conditions), surface scattering becomes the limiting factor in the double-gate device and the mobility decreases linearly with field in the log-log scale. Still, the slope of the linear decrease in the log-log scale is slightly lower

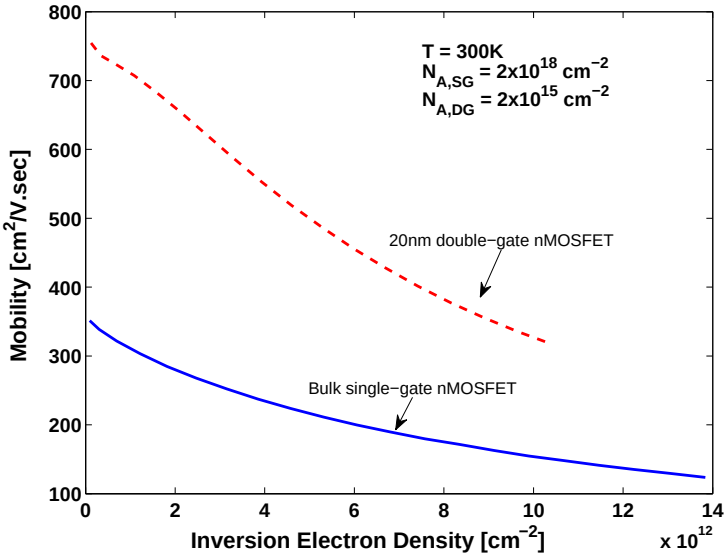


Figure 6-6. Electron mobility vs. inversion electron density is shown for a bulk single-gate (100) MOSFET with doping density of $N_{A,SG} = 2x10^{18}cm^{-2}$ and for a 20nm thick double-gate (100) MOSFET with doping density of $N_{A,DG} = 2x10^{15}cm^{-2}$ at 300K.

for the double-gate device compared to the bulk single-gate device, which is explained by the non-zero density of inversion electrons deep inside the fin away from the rough surface.

The single-gate MOSFET effective low-field mobility is simulated for a temperature range of 400K to 100K with decrements of 50K and is shown in Figure 6-8. At elevated temperatures, independent of the vertical electric field, the increase in phonon scattering degrades the electron mobility significantly. The total phonon limited mobility increases from $\sim 500cm^2/V.sec$ to around $\sim 4000cm^2/V.sec$ as the temperature decreases from 400K to 100K. Hence, the surface roughness scattering limited mobility becomes the dominant scattering mechanism for low temperatures. The 100K surface roughness limited mobility at low gate bias ($E_{eff} = 0.8MV/cm$) is around $420cm^2/V.sec$ and sets the level of the total mobility at this temperature. Since the carriers spread deep into the bulk at high temperatures, the surface roughness limited mobility at low gate bias increases $\sim 60\%$ to $690cm^2/V.sec$ at 400K. However at high effective vertical fields, the

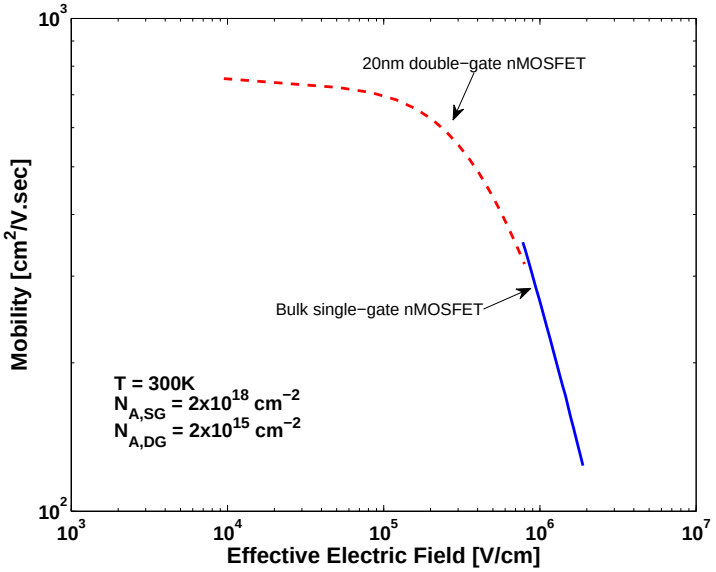


Figure 6-7. Electron mobility vs. effective electric field is shown for a bulk single-gate (100) MOSFET with doping density of $N_{A,SG} = 2x10^{18} cm^{-2}$ and for a 20nm thick double-gate (100) MOSFET with doping density of $N_{A,DG} = 2x10^{15} cm^{-2}$ at 300K.

temperature dependence of surface roughness scattering is negated since the strong electric field brings electrons close to the interface at all temperatures. Roughly around $E_{eff} = 1.4 MV/cm$, the surface roughness limited mobility becomes almost the same for all temperatures. Consequently at high gate bias (higher fields), the surface roughness limited mobility sets the limit for the total mobility and the temperature dependence of the phonon limited mobility defines the difference between the total mobility at different temperatures.

The temperature dependence of the inversion electron density and the normalized linear drain current vs. gate voltage are simulated and shown in Figure 6-9 and Figure 6-10. For the same doping density and gate workfunction, the threshold voltage shift due to temperature is found to be around $-0.6 mV/K$, which is consistent with the literature [156]. The normalized linear drain current at $V_{DS} = 0.05 V$ is calculated by,

$$\frac{I_{lin}}{W} = \frac{\mu q N_{inv} V_{DS}}{L}. \tag{6–20}$$

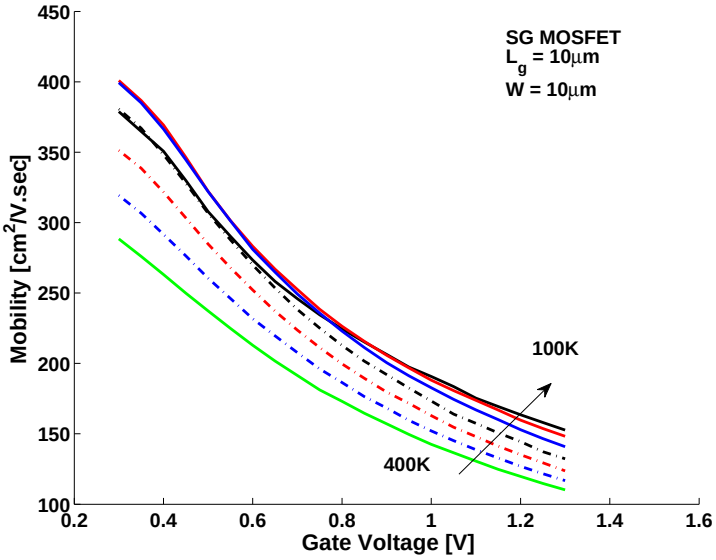


Figure 6-8. Simulated single-gate MOSFET low-field electron mobility vs. gate voltage is shown for a temperature range of 400K to 100K.

As expected, at low temperatures, the threshold voltage is shifted to higher values, and the linear drain current exhibits a steeper increase compared to that at elevated temperatures.

The normalized strain-induced change in linear drain current is simulated at 300K for a uniaxial stress range of 75MPa tension to 75MPa compression, without the inclusion of an *ad hoc* modification of the surface roughness parameters. The hydrostatic and uniaxial deformation potentials for strain effects on conduction band structure are taken from the pseudo-potential calculations of [140]. In [140], the shear strain deformation potential is empirically chosen to match the nMOSFET piezoresistance coefficient, which may overshadow the possible effects of surface roughness morphology modification and attribute the enhanced mobility change to a larger band warping induced effective mass reduction. Thus, we have modeled the band warping induced mass reduction using the experimental results of Hensel *et al.* [144], as shown in Eq. 6–6. The simulated change in channel resistance at $E_{eff} = 1MV/cm$ is shown in Figure 6-11, along with the expected change calculated

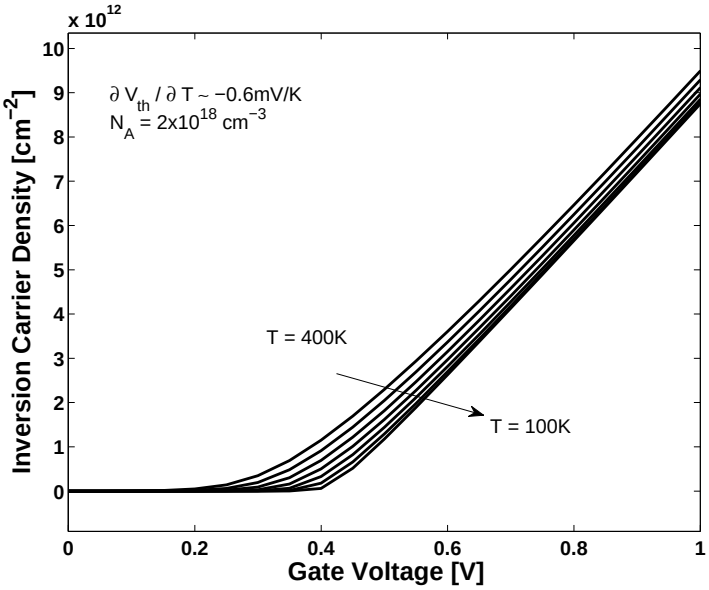


Figure 6-9. For the single gate MOSFET, simulated inversion carrier density vs. gate voltage is shown for temperatures ranging from 400K to 100K with decrements of 50K. The threshold voltage shift due to temperature is found to be around -0.6mV/K.

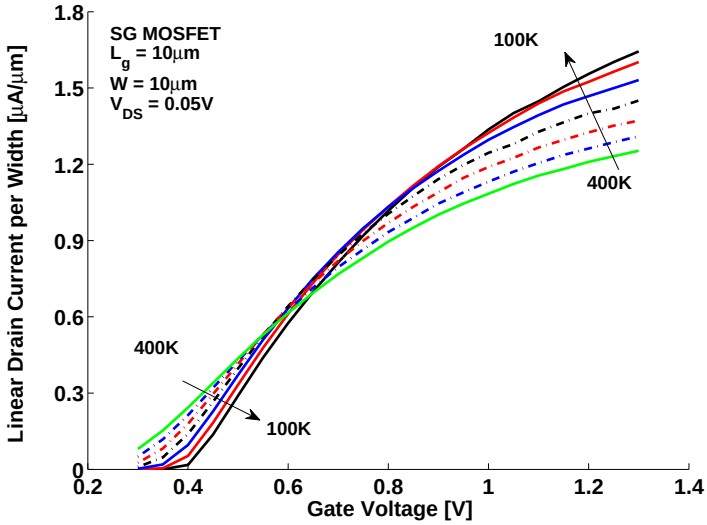


Figure 6-10. Simulated single-gate MOSFET linear drain current vs. gate voltage is shown for a temperature range of 400K to 100K.

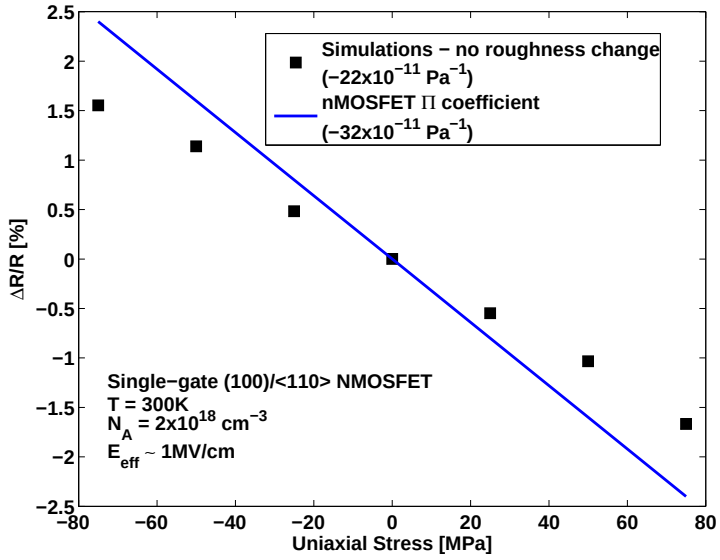


Figure 6-11. Simulated single-gate nMOSFET normalized channel resistance change vs. uniaxial stress (symbols) and the expected change from nMOSFET π-coefficient (line) are shown for $E_{eff} = 1MV/cm$. For the simulation, *ad hoc* modification of surface morphology is not included. The simulated Π-coefficient ($-22 \cdot 10^{-11} Pa^{-1}$) is smaller than expected ($-32 \cdot 10^{-11} Pa^{-1}$).

by the nMOSFET π-coefficient [104]. Without an *ad hoc* modification of the surface morphology with strain, the simulated π-coefficient ($-22 \cdot 10^{-11} Pa^{-1}$) is found to be smaller than the expected value of $-32 \cdot 10^{-11} Pa^{-1}$. Consequently, a modification of the surface roughness parameters under mechanical stress is necessary to understand the governing physics of strain-induced enhancement of electron mobility in silicon inversion layers.

To match the experimentally extracted piezoresistance coefficient, we have simulated a large matrix of Δ and L pairs for three stress conditions (25, 50 and 75MPa). Once a close agreement is obtained for a given stress condition, the *ad hoc* changes in the surface morphology is linearly interpolated to other stress conditions and the π-coefficients are calculated under these conditions. For simulations of all three stress conditions, a 1.7% decrease in Δ and a 1.7% increase in L per 100MPa uniaxial tension resulted in a constant π-coefficient of $-32 \cdot 10^{-11} Pa^{-1}$. Figure 6-12 shows the details

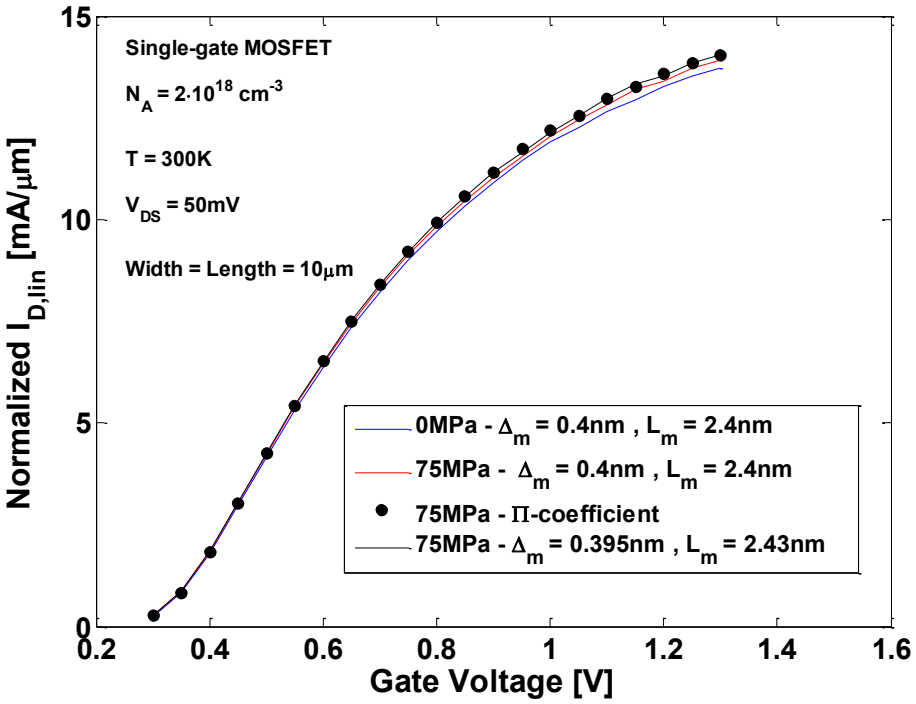


Figure 6-12. The simulation results are shown for the 75MPa condition. For the simulated unstrained linear drain current shown with blue line, the expected drain current at 75MPa using the experimental π-coefficient is shown with symbols. The simulated current at 75MPa with and without an *ad hoc* strain-induced surface morphology are shown with black and red lines.

of the simulation results for the 75MPa condition. The unstrained linear drain current (blue line) and the expected 75MPa drain current using the experimental π-coefficient (symbols) are shown. Simulated current-voltage curves at 75MPa with and without an *ad hoc* surface morphology change are shown with black and red curves. While the absence of the surface modification leads to an underestimated drain current at 75MPa, the agreement between the expected and simulated current is very close using a modified set of Δ and L parameters. The normalized error in the *ad hoc* fitting is shown in Figure 6-13 for all three stress conditions, using a -1.7%/100MPa change in Δ and +1.7%/100MPa change in L. The error in all stress conditions is less than 0.05% at the strong inversion operation.

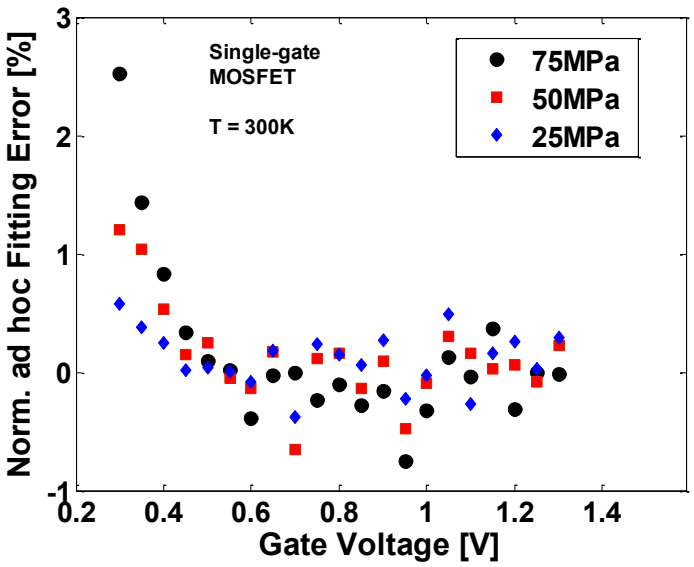


Figure 6-13. Normalized *ad hoc* fitting error is shown for 25MPa, 50MPa and 75MPa conditions for a given gate voltage.

The fitted value of a 1.7% magnitude change in both Δ and L per 100MPa uniaxial tension is quite large. In fact, it is almost equivalent to the silicon Young's modulus lowered to 6GPa. However, the AFM measurements carried out in our group also indicate an equivalent Δ reduction of 3.4% for 200MPa uniaxial tension [138]. Recently, a similar AFM data on the roughness of silicon surfaces with uniaxial tension is published, claiming a 1.9% decrease in Δ per 100MPa of uniaxial tension [157]. Although these independent results are consistent, the suggested change in the lattice for the given stress levels is enormous. However, it has been shown through molecular dynamics simulations that the Young's modulus of silicon nanowires decreases significantly at very small thicknesses due to reduction of crystal rigidity [58]. Therefore, a similar local reduction in the silicon Young's modulus within a couple of monolayers near the oxide boundary could be the reason for the observed strain-induced surface morphology change.

6.5 Summary

The experimental evidence from electrical and materials characterization suggests reduced interface roughness at the Si/SiO_2 boundary. By using a self-consistent

Schrödinger-Poisson simulator and an *ad hoc* assumption of strain-induced surface morphology change, we have shown that the improved electron transport seen at high inversion carrier densities and tensile stress is likely to be a result of a smoother Si/SiO_2 interface. While the claimed change in the surface morphology is quite large for the given amount of mechanical stress, consistent results are observed from both our group and other researchers in the literature. A possible lower Young's Modulus at the silicon crystal surface is expected to be the reason for the enhanced strain-induced morphology change, and thus the carrier transport at high inversion carrier densities.

CHAPTER 7
UNIFIED QUASI-BALLISTIC COMPACT MODEL FOR STRAINED NANOSCALE MOSFETS USING AN UPDATED ONE-FLUX THEORY

7.1 Motivation

The incessant pursuit of Moore's law [158][159] and transistor scaling over the past four decades has led the semiconductor industry towards high performance transistor architectures at sub-100 nm channel lengths today [160]. As device dimensions approach technological and fundamental limits [161], a solid understanding of the carrier transport mechanism is of paramount importance. Recently, there has been significant interest in re-evaluating and re-examining the traditional carrier transport models in MOSFETs towards capturing and explaining off-equilibrium transport phenomena like velocity overshoot [162][163] observed experimentally. It is of interest to note that this effort has been ongoing at all levels of modeling–numerical, empirical, compact and physics-based. The conflicting requirements of accuracy, generality and computational accuracy only makes this task more difficult–especially with shrinking channel lengths, where the carrier transport is thought to become more and more ballistic.

Carrier transport in semiconductor devices is analyzed using a wide spectrum of approaches. These range from the macroscopic Drift-Diffusion (DD) [164] approach, to the semi-classical Hydrodynamic (HD) and energy-transport models [165] employing higher moments of the Boltzmann Transport Equation (BTE), to the very rigorous numerical solutions of the BTE using Monte Carlo techniques [166][167]. However, in the last decade, methods based on McKelvey/Shockley's alternative one-flux approach [168][169][170][171] have emerged as a simplified alternative for understanding the physics of transport in nanoscale devices. Especially, the one-flux theory based Lundstrom model (LM) [172] has found a lot of traction owing to its simple expressions giving calculated currents close to those predicted by Monte carlo methods.

Both the DD and one-flux methods have been used well beyond their nominal range of applicability, and the reason for this success was investigated by Lundstrom

and co-workers [173]. They reported that the one-flux theory is still applicable in analyzing transport, even though the assumption of carriers moving at equilibrium thermal velocity need not necessarily hold. They attributed this to the presence of a the current-limiting “bottleneck” in a transistor (controlled by a gate or a base) and that this bottleneck region is present in a low-field region adequately described by the DD/one-flux approaches. The LM adopts a transmission picture of the transistor. In this picture, carriers are injected from a thermal equilibrium source reservoir, across a potential energy barrier , into the channel and finally exit through the drain. The height of the potential barrier is modulated by the applied voltage on gate that is electrostatically coupled to the channel. Though current flow is dependent on the voltage applied on the drain, the density of carriers in the channel are determined by the device electrostatics that come into play to maintain charge balance.

Using the one-flux theory approach, the LM makes it possible to quantify device performance in terms of two key parameters namely, the ballistic efficiency B (a measure of how the current in a given device compares to that of an ideal, scattering free ballistic transistor of the same dimensions) and the carrier injection velocity v_{inj} (the average velocity of the carriers at the top of the source-channel potential barrier in the same device). Several numerical [174][175][176][177][178] and experimental studies [179][180][181][182][183][184][185][186][187][188][189] that support the overall trends discussed in the LM have been published. Some authors have also questioned [190][191] the validity of the LM, pointing out a few underlying theoretical assumptions. Recently, Natori proposed a new solution [192][193] for understanding the high-field transport in bulk semiconductors, that could be applicable for understanding quasi-ballistic transport.

Since its introduction in 2004 [194], uniaxial stress has been a proven industry technique for enhancing transistor performance. The benefits of uniaxial stress continues to be exploited for current technologies [195] and is projected to be an

important factor even at sub-10 nm channel lengths [196]. While the effects of stress on carrier mobility has been well studied, understanding its impact on quasi-ballistic transport is listed as one of the challenges the ITRS [197] for emerging devices. There have also been several experimental works published, for example [198][199] [200][201], that attempt to extend the LM towards explaining the experimentally observed results on the ballistic efficiency of strained devices. Most of these works simply extend the LM empirically to explain the observed effects, and it is difficult to get a clear understanding of the strain effects. This is because of three reasons (1)fundamental assumptions within the LM framework for high field transport (2) inherent effect of strain on carrier mass and carrier scattering (3) coupled nature of carrier transport and device electrostatics.

The purpose of this chapter is twofold: (1) A thorough review and analysis of the underlying assumptions and shortcomings in the LM and LM based methods for evaluating the ballistic efficiency of a nanoscale MOSFET (2) Presenting an updated, surface-potential based, unified model of transmission coefficient using an energy averaged one-flux theory. This transmission coefficient could to be used to better study the effect of strain in comparison to existing methods.

7.2 Fundamentals

7.2.1 Natori's Original Ballistic Transport Model

In 1994, Natori [202] first proposed and studied transport in a completely ballistic MOSFET. a thin Si film connected to two reservoirs of carriers (source and drain) on either side. The gate electrode is electrically separated from the Si film by a dielectric. The Si film is left undoped to reduced scattering in the channel. In the transport model described, the MOSFET I–V relation is expressed in terms elementary parameters without any dependence on the carrier mobility because of the absence of scattering. The current is independent of the channel length and is proportional to the channel

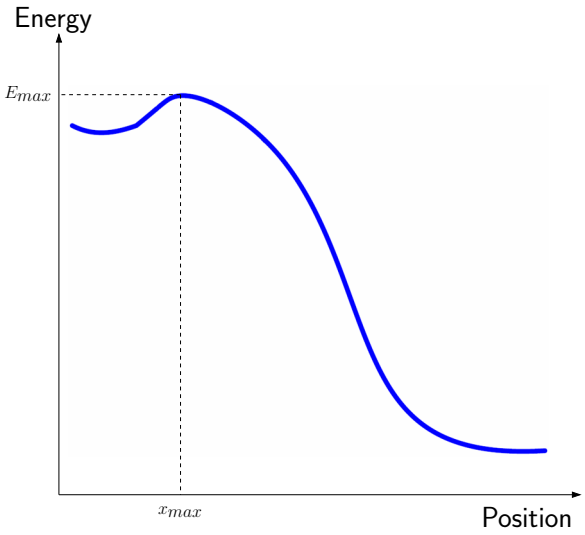


Figure 7-1. Schematic of conduction or first subband profile along the length of a MOSFET channel for device biased in saturation conditions.

width. The current value saturates as the drain voltage is increased and the linear and saturation operations are specified as in a conventional MOSFET.

For purely ballistic transport, modeling the transport involves the study of carrier transmission over the source-to-drain potential barrier (Fig.(7-1)). Electrons possessing energies higher than the barrier height E_max are transmitted from source reservoir into the channel by thermionic emission. Similar phenomenon happens at the drain reservoir–however, once a bias is applied at the drain, the barrier height for the drain injected carriers becomes really high as shown in the figure, and drain injection is suppressed. The starting assumptions made by Natori are:

1. Channel potential varies gradually along the channel length L from a high value at the source to a low value at the drain. A maximum value E_{max} occurs at $x = x_{max}$.

2. The carriers are confined along the channel width W by the steep barrier at the edges. The potential profile along the width is approximated by a square potential well for channel widths larger than the range of potential variation.

3. Carrier confinement (y-direction) takes place due to a sharp triangular potential well with the electronic states expressed by discrete energy levels, similar to a regular MOSFET.

Natori also derived a compact equation for the $I-V$ relation:

$$I_D = W \cdot \frac{\sqrt{2}q(K_B T)^{3/2} M_v m_t^*}{\pi^2 \hbar^2} \left\{ \mathcal{F}_{1/2}(u) - \mathcal{F}_{1/2}(u - v_D) \right\} \tag{7–1}$$

where m_t^* is the transverse effective mass of the electron, M_v = (lowest valley degeneracy) $\times$ (total population/population of lowest levels). $\mathcal{F}_{1/2}$ is the Fermi-Dirac integral of order 1/2 defined as:

$$\mathcal{F}_{1/2}(x) = \int_0^\infty \frac{\sqrt{x}\,dy}{1 + \exp(y - x)} \tag{7–2}$$

The other parameters in Eq.(7–1) are defined as below:

$$\begin{aligned} u &= \ln[((1 + \exp(v_D))^2 + 4 \cdot \exp(v_D) \cdot \exp(\rho - 1))^{0.5} \\ &\quad - (1 + \exp(v_D))] - \ln(2) \\ v_D &= \frac{qV_D}{KT} \\ \rho &= \frac{2\pi\hbar^2}{qK_B T m_t^* M_v} \cdot (C_G(V_G - V_{TH})) \end{aligned} \tag{7–3}$$

Since this model describes purely ballistic transport with no scattering of any kind inside the channel, the resulting current is the maximum current obtainable for a particular geometry. It is well known that transport in sub-50nm MOSFETs are quasi-ballistic and Natori's original model, therefore, is not directly applicable for describing transport where only a fraction of the carriers reach the drain without undergoing scattering.

7.2.2 Current Control in Natori's Model

In conventional MOSFET theory, the drain current is governed by the carrier velocity in the channel. This carrier velocity is expressed as a product of the lateral electric field and the carrier mobility when the electric field is not as large. When the electric field is increased, however, the carrier velocity saturates to a value around 10^7 cm/s due to energy relaxation via optical phonon scattering. In the fully ballistic MOSFETs in Natori's discussion , carriers in the channel are free from scattering and all carriers propagating

towards the drain, passing through the bottleneck around x_{max}, reach the drain without scattering back to the source, if the reflection at the drain edge is neglected. The channel current is governed by the current at the bottleneck. In current saturation, the number of carriers injected from the source to the bottleneck dominates the total current, and the carrier velocity in the backward channel has little to do with the current value. For degenerate carriers, the mean carrier velocity (known as the injection velocity) at the bottleneck is expressed as:

$$v_{inj} = \frac{8\hbar\sqrt{Q}}{3m_t^*\sqrt{q\pi M_v}} \simeq \frac{8\hbar\sqrt{C_G(V_G - V_{TH})}}{3m_t^*\sqrt{q\pi M_v}} \tag{7–4}$$

To quantify the ballistic transport in devices, Natori also came up with the Ballistic Efficiency parameter B that is defined as the ratio of the experimental $I_{D(\text{sat})}$ to the theoretical ballistic $I_{D(\text{sat(bal)})}$. Since the current depends on the injection velocity, we have

$$B = \frac{v_{inj}}{v_{inj(bal)}} = \frac{I_{D(\text{sat})}}{I_{D(\text{sat(bal)})}} \tag{7–5}$$

Natori predicted that B will increase as the carrier scattering is reduced and will eventually reach unity, theoretically, when ballistic transport is realized.

7.2.3 One-flux Theory: Basics

Following McKelvey's approach [168][169][171], the basic formalism of the one-flux theory can be understood by considering Figure 7-2, where the position (and time) dependent carrier fluxes are shown at a slab of thickness dx for a positively charged carrier. The channel of a transistor can be thought of a cascade of many such slabs, with the output of one slab becoming the input of the next in a continuous fashion. For simplicity, we ignore both the generation/recombination processes inside the device and the time dependence of the various parameters. The slab itself can be characterized by a scattering matrix, which links the incoming and outgoing fluxes as

$$\begin{bmatrix} J^{+}(x+dx) \\ J^{-}(x) \end{bmatrix} = \begin{bmatrix} t^{+}(x) & r^{-}(x) \\ r^{+}(x) & t^{-}(x) \end{bmatrix} \begin{bmatrix} J^{+}(x) \\ J^{-}(x+dx) \end{bmatrix}$$

Here $r^{\pm}(x)$ and $t^{\pm}(x)$ represent the ratio of carriers reflected and transmitted in a given direction. If the expressions for these coefficients (in terms of position and field) and the field profile in channel the are known, it is possible to calculate the scattering matrix for the entire channel. From that matrix, the overall transmission coefficient $\mathbf{t}$ and reflection coefficients $\mathbf{r}$ of the entire device can be calculated for a family of gate/drain voltages applied to the device. Since we ignore generation and recombination, it follows that $\mathbf{t} + \mathbf{r} = 1$.

This method is called the one-flux method because only one carrier flux is considered in each direction. In reality, since carriers entering the channel have a distribution of incoming energies, each energy level will have its own flux components and there would be intermixing of the fluxes between different energy levels after momentum altering scattering events. Theoretically, it is possible to do a energy summation of the fluxes to arrive at an overall one-flux equation. Assad and Lundstrom [173] formulated a M-flux theory that reformulated the BTE as M coupled continuity and DD equations in the energy-momentum space. As would be expected, the details of transport over the barrier using this approach is entirely numerical and while accurate, it is also very complex to implement. It was resolved in their work that the simple analytical expressions of the one-flux theory (discussed below) fail in the high field conditions, and that they should be used for only first-order calculation purposes. However, owing to the simplicity of approach, the quasi-ballistic transport model formulated later by Lundstrom for understanding the essential physics of carrier transport [203] utilizes the one-flux theory assumptions and approximations. Since our goal is to improve upon the one-flux theory towards a better understanding of strain effects with a more accurate analytical

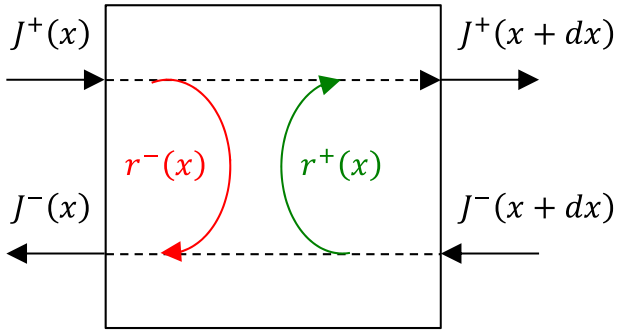


Figure 7-2. Incoming and outgoing carrier fluxes in a slab of thickness dx. $r^+(x)$ and $r^-(x)$ are the reflected fluxes in the forward (parallel to applied field) and reverse (antiparallel to applied field) directions respectively. Carrier generation/recombination within the slab is ignored for simplicity.

expression for the transmission coefficient, we focus on the one-flux theory approach in this chapter.

The differential equations for the flux method, ignoring recombination/generation, are as below (detailed derivation can be found in [168][170]).

$$\frac{dJ^+}{dx} = -r^+J^+ + r^-J^- = \frac{dJ^-}{dx} \tag{7–6}$$

Here J^+ and J^- denote the postion/field dependent positive(left-directed) and negative (right-directed) fluxes. r^+ and r^- are the postion/field dependent backscattering coefficients in the appropriate directions.

The net flux flowing left can be written as

$$J = J^+ - J^- \tag{7–7}$$

We have dropped the subscripts for simplicity – i.e. $J^+(x, E_x)$ is simply written as J^+. E_x denotes the electric field in the left direction.

The effective carrier concentration is given by

$$n^* = \frac{J^+}{v^+} + \frac{J^-}{v^-} \tag{7–8}$$

where v^+ and v^- are the velocities of the two fluxes. Assuming that each of these is equal to the equilibrium 1-D carrier velocity of non-degenerate semiconductor with effective mass m^*, given by

$$c = \sqrt{\frac{2k_B T}{\pi m^*}} \tag{7–9}$$

it is very easy to show that

$$2J^{\pm} = cn^* \pm J \tag{7–10}$$

Using above equations, we can derive an expression for the net flux J as

$$J = -\frac{r^+ - r^-}{r^+ + r^-} cn^* - \frac{c}{r^+ + r^-} \frac{dn^*}{dx} \tag{7–11}$$

For small electric fields, it is assumed that the $(r^+ - r^-)$ should be proportional to the electric field (near equilibrium conditions), and that $(r^+ - r^-)$ be independent of the field,

$$r^+ - r^- = \eta E_x \tag{7–12}$$

and therefore, we can write Eq. 7–11 as

$$J = -\frac{c\eta}{r^+ + r^-} n^* E_x - \frac{c}{r^+ + r^-} \frac{dn^*}{dx} \tag{7–13}$$

This equation can be written in a familiar form similar to the Drift-Diffusion equation, if we define mobility-like and diffusion-constant-like terms μ^* and D^*

$$\mu^*(x, E_x) = \frac{c\eta}{r^+ + r^-} \tag{7–14a}$$

$$D^*(x, E_x) = \frac{c}{r^+ + r^-} \tag{7–14b}$$

$$\frac{D^*}{\mu^*} = \frac{1}{\eta} \tag{7–14c}$$

to arrive at

$$J = -\mu^* n^* E_x - D^* \frac{dn^*}{dx} \tag{7–15}$$

The above equivalence between the DD theory and the one-flux equations was discovered by Shockley in 1962 [170], where he also observed that the diffusion constant D^* of the carriers traversing a infinitesimally small field-free region (one that is smaller than the average mean free path) is the same as that of the bulk semiconductor, D_0.

If the one flux equations have to be compatible within the framework of general semiconductor transport equations, it is clear that

$$\lim_{E_x \to 0} \eta = \eta_0 = \frac{q}{k_B T} \tag{7–16a}$$

$$\lim_{E_x \to 0} \mu^* = \mu_0 = \eta_0 \frac{c}{2r_0} = \frac{q}{k_B T} \frac{1}{D_0} \tag{7–16b}$$

where μ_0, k_0, D_0 denote equilibrium quantities.

McKelvey showed ([168], Appendix B) that

$$r_0 = \lim_{E_x \to 0} r^{\pm} = \frac{3}{4\lambda_s} \tag{7–17}$$

Here λ_s denotes the "scattering mean free path". Note that Shockley argues ([170], Appendix 1) that McKelvey's expression for r_0 needs a small correction.

7.2.4 Revisiting the Lundstrom Model

Lundstrom used the one-flux theory to arrive at a compact expression for the drain current in a MOSFET. We start this section with a discussion on the concept of ballistic mobility which explains the link between the one-flux theory and Lundstrom's model.

The phenomenological effective mobility that is typically used in the MOSFET $I - V$ equations was thought to be independent of channel length. However, experimental measurement of this μ_{eff} from short channel MOSFETs is seen to be considerably smaller than the long channel mobility μ_0. Shur [204] first linked this degradation to the finite carrier acceleration time in the channel. To account for this, he introduced a

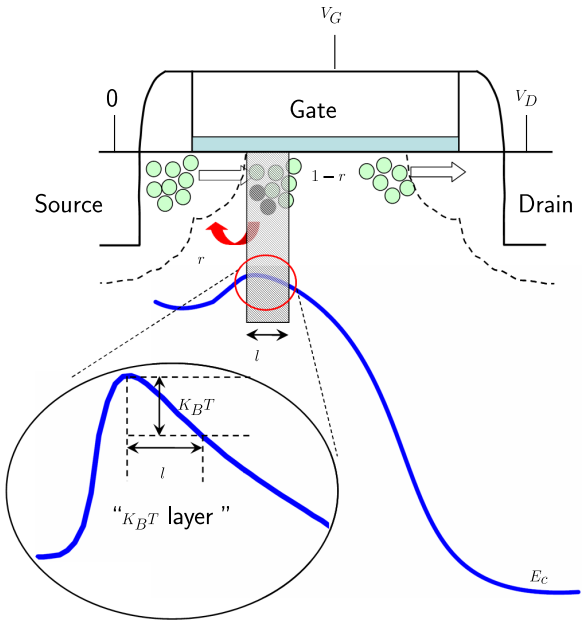


Figure 7-3. Definition of the KT-layer in the Lundstrom Model, showing transmitted and backscattered carriers near the critical region for transport.

"ballistic" mobility μ_B, which he defined as:

$$\mu_B = \frac{qL}{m_c^* \pi v_T} \tag{7–18}$$

where L is the channel length, m_c^* is the conductivity effective mass and $v_T = \sqrt{2k_B T/\pi m_c^*}$ is the thermionic emission velocity (same as the 1-D velocity c described in the last section). The short channel mobility was then formulated in a Mattheissen's rule form as:

$$\frac{1}{\mu_{\text{eff}}} = \frac{1}{\mu_B} + \frac{1}{\mu_0} \tag{7–19}$$

Shur noted that the finite acceleration time L/v_T in Eq. (7–18) describes the time required for a carrier to travel over a length L with the finite velocity v_T given at the beginning of the channel. Shur's final $I-V$ expression is simply an extension of the classic DD-based MOSFET equation.

$$I_D \simeq \frac{W}{L} C_G \left(\frac{1}{\mu_B} + \frac{1}{\mu_0}\right)(V_G - V_{TH})V_D \tag{7–20}$$

When the ratio μ_B/μ_0 is small, i.e for long channel devices where the transport is not ballistic at all, the above equation reduces to the traditional DD equation.

In the landmark paper [205] written in 1997, Lundstrom gave a re-interpretation of Shur's Eq. (7–20) for a quasi-ballistic MOSFET operating in saturation conditions. Following McKelvey's Flux-based approach, the MOSFET is treated as a cascade of three regions: source-channel-drain. The source is a reservoir of thermal carriers which injects a flux of carriers into the channel over a potential barrier. The gate modulates the height of this potential barrier. Near the barrier, a fraction r of the source flux backscatters from the channel and re-enters the source. The remaining fraction $1-r$ transmits and eventually enters the drain. This is conceptually shown in Fig.(7.2.4).

It was shown in [205] that $I_{D(\text{sat})}$ can be compactly expressed as:

$$I_{D(\text{sat})} = C_{\text{ox}}(V_G - V_{TH})\frac{1-r_{sat}}{1+r_{sat}}v_T \tag{7–21}$$

where r_{sat} is the ratio of the backscattered carriers (i.e 'backscattering ratio') under saturation conditions. v_t is the velocity of the forward going electrons (assuming nMOSFET) at the the top of the source-channel potential barrier (i.e the 'virtual source'). The backscattered electrons are assumed to have this velocity as well. This velocity is simply the 1D thermal velocity in the transport direction, given by $\sqrt{\frac{2K_BT}{\pi m_c^*}}$. The ratio $\frac{1-r_{sat}}{1+r_{sat}}$ is referred to as the Ballistic Efficiency, B. The product of B and v_T is called the injection velocity. Note that this definition is slightly different from Natori's definition.

The appeal of Lundstrom's theory stems from the simplicity of Eq. (7–21) that describes the saturation current. From McKelvey's scattering theory [168][170], the

---

$$\frac{I_{D(\text{sat})}}{W} = Q_i(0)\left(\frac{1-r}{1+r}\right)\left\{v_T\frac{\mathcal{F}_{1/2}(\eta_F)}{\mathcal{F}_0(\eta_F)}\right\}\left\{\frac{1-\dfrac{\mathcal{F}_{1/2}(\eta_F-U_D)}{\mathcal{F}_{1/2}(\eta_F)}}{1+\left(\dfrac{1-r}{1+r}\right)\dfrac{\mathcal{F}_0(\eta_F-U_D)}{\mathcal{F}_0(\eta_F)}}\right\} \tag{7–33}$$

backscattering ratio is actually well defined for zero (or very small) electric fields as:

$$r_0 = \frac{L}{L+\lambda} \tag{7–22}$$

where L is the channel length λ is the mean- free path (or more appropriately the velocity/momentum relaxation length) of the electrons just near the virtual source. In the non-degenerate limit, the parameter λ is actually related to the low field, long channel mobility as:

$$\lambda = 2\frac{K_B T}{q}\frac{\mu_0}{v_T} \tag{7–23}$$

Based on an earlier observation by Price [206] that an electron which has traveled "a few times 10^{-6} cm" in a field of 10^4 V/cm, so has "descended farther than this has little chance of getting back all the way" and will "almost certainly reach the [drain]", (i.e. electrons that have traveled past a critical distance from the virtual source cannot return back to the source) Lundstrom speculated that the backscattering ratio in high field conditions can be defined as:

$$r_E = \frac{l}{l+\lambda} \tag{7–24}$$

Here l is some critical dimension which is assumed to be the width of the so-called $K_B T$-layer in [205].To first order, $l \approx K_B T/q \cdot V_D/L$. The width of the $K_B T$-layer, as shown in Fig.(7.2.4), is simply the distance from the virtual source over which the potential drops by $K_B T$. Thus,

$$r_E = \frac{1}{1+\frac{\lambda}{l}} = \frac{1}{1+\frac{2\mu_0}{v_T}\frac{V_D}{L}} \tag{7–25}$$

Combining Eqs. (7–23) and (7–24) , a more general expression for the backscattering ratio in saturation can be written as

$$r_{sat} = \frac{r_0}{1+\frac{2\mu_0}{v_T}\frac{V_D}{L}} \tag{7–26}$$

To establish a connection between this model and conventional DD-based models, consider a device with channel length L long enough so that $r_0 \simeq 1$ (i.e $r_{sat} \simeq r_E$). Then, using Eqs. (7–21) and (7–26), Lundstrom came up with the following expression for the quasi-ballistic linear MOSFET $I-V$ equation:

$$I_{D(\text{sat})} = WC_G \left\{ \frac{\frac{1}{rsat} - 1}{\frac{1}{rsat} + 1} \right\} v_T (V_G - V_{TH}) \tag{7–27a}$$

$$= WC_G \left\{ \frac{1}{\frac{1}{v_T} + \frac{L}{\mu_0 V_D}} \right\} (V_G - V_{TH}) \tag{7–27b}$$

Compare this equation with Shur's Equation (7–20), which can be rewritten as:

$$I_D \simeq WC_G \frac{1}{\frac{L}{V_D}\left(\frac{1}{\mu_B} + \frac{1}{\mu_0}\right)} (V_G - V_{TH}) \tag{7–28}$$

This shows the link between traditional models and Lundstrom's $K_B T$-layer model. In [172][207], Lundstrom and Ren further clarified their model explaining the scattering mechanism and current control mechanism in more detail with detailed simulations based on the Non-equilibrium Green's Function formalism. In this work, carrier backscattering in the KT-layer was quantified and shown to be crucial in estimating how close a given device was close to the ballistic limit. We briefly discuss some aspects of the physics involved in the next section.

7.2.5 Current Control in Lundstrom Model

Lundstrom pointed out in [207] that that the maximum average carrier velocity at the beginning of the channel was the equilibrium, uni-directional thermal velocity (as first pointed out by Natori). For degenerate conditions, this velocity is:

$$v_T(V_D) = \sqrt{\frac{2K_B T}{\pi m_c^*}} \left\{ \frac{\mathcal{F}_{1/2}(\eta_F)}{\mathcal{F}_0(\eta_F)} \right\} \tag{7–29}$$

where $\eta_F = \dfrac{E_F - E_1}{K_B T}$ with E_F being the Fermi energy and E_1 being the energy level of the lowest subband. For simplicity, it is assumed that most of the carriers reside in the first subband.

The complete equation for $I_{D(\text{sat})}$ for the full range of bias conditions on the gate and the drain terminals of a MOSFET is shown at the bottom of the page (Eq. (7–33)). Here,

a. The first term $Q_i(0)$ is the inversion charge at the virtual source. For a well designed MOSFET, $Q_i(0) \simeq C_G(V_G - V_{TH})$ above threshold.

b. The second term is the ballistic efficiency B, which describes the reduction of $I_{D(\text{sat})}$ due to backscattering. The parameter r is a function of both V_G and V_D.

c. The third factor is the degenerate thermal velocity which depends on $Q_i(0)$ through η_F.

d. The last factor accounts for the dependance on the drain bias V_D. This factor is proportional to $U_D = \dfrac{qV_D}{K_B T}$ for low V_D and approaches unity for high V_D conditions, thereby reducing to the form of Eq.(7–21). Also note that when $r = 0$, this equation reduces to Natori's ballistic MOSFET equation.

The backscattering ratio r (and ultimately $I_{D(\text{sat})}$) depends on the mean free path (λ) in the critical region and the width (l) of the critical region itself . λ, in turn depends on the low field, long channel mobility. While Eq. (7–23) is applicable in non-degenerate conditions, the complete expression for λ is [207]

$$\lambda = 2\frac{K_B T}{q}\frac{\mu_0}{v_T}\frac{\mathcal{F}_0(\eta_F)}{\mathcal{F}_{-1}(\eta_F)} \tag{7–30}$$

The width of the $K_B T$-layer was modeled as

$$l = L\left(\kappa\frac{K_B T}{qV_D}\right)^{\beta_l} \tag{7–31}$$

where κ and β_l are some empirical constants. The gate voltage dependence of the layer comes through the β_l parameter. The original LM assumed $\beta_l \approx 0.7$ and $\kappa > 1$.

Thus it was argued that the *long channel mobility* in a given technology is still an important parameter even for nanometric MOSFETs in the same technology, that are

presumed to be quasi-ballistic. Lundstrom also showed that the velocity saturation occurs in a ballistic MOSFET–but agaist the conventional picture of saturation through high field optical phonon emission–it was indicated that the velocity at the virtual source (i.e the top of the barrier) saturates at the thermal limit. For unstrained Si, these two velocities, i.e. the field limited saturated velocity (v_{sat}) and the maximum uni-directional thermal velocity ($v_{inj} = B \cdot v_T$, with $B = 1$) both happen to be really close to each other ($\sim 10^7$ cm/s).

Lundstom also argued that the backscattering that happens in the $K_B T$-layer is the most critical scattering in the MOSFET. According to [172], only this backscattering determines the ballistic efficiency and the $I_{D(\mathrm{sat})}$. The reasoning given is quite simple: even if backscattering occurs beyond the critical distance l, the carriers will not have enough *longitudinal* energy to surmount the barrier and re-enter the source.More likely, the carriers will undergo several scattering events (phonon/impurity), which will further reduce the longitudinal energy, and they will eventually reach the drain.

7.2.6 Issues in the Lundstrom Model

Advanced Monte Carlo simulation performed by Palestri [175] in 2005 brought out some important issues in the $K_B T$-layer theory. Some of these issues stem from the fundamental assumption on the nature of backscattering ratio (esp. in high V_D) and its dependence on the low field, long channel mobility. Meanwhile, based on a temperature-based experimental method pioneered by Chen [179] in 2002, several authors [176, 180, 181, 183, 184, 186–189, 199, 201, 208–212] published have widely varying numbers and interpretations for the ballistic efficiency B in a modern MOSFET, both with and without strain.

To simplify the discussion, we follow the terminology used by Palestri [175]. At the virtual source, the current due to backscattered carriers is denoted by I_{on}^{-} and the current due to forward moving carriers is denoted by I_{on}^{+}. The backscattering ratio r is simply $\frac{I_{on}^{-}}{I_{on}^{+}}$. If we call the total current reaching the drain in the case of complete ballistic transport as

I_{bal} and in the presence of "scattering" as I_{on}, it follows that

$$I_{on} = \left(\frac{1-r}{1+r}\right) I_{bal} \tag{7–32}$$

One way of interpreting this equation is as follows: if we can calculate the complete ballistic current using say a device simulator, then for a real device which has scattering, all we need is to determine r to estimate the on current. However, the extraction of precise numbers for r is clouded by a number of theoretical and experimental uncertainties.

The following is a list of implicit and explicit assumptions in the $K_B T$-layer theory, including the applicability and the validity of the assumptions.

1. At the top of the barrier (virtual source), there are two streams of carriers (Fig.(7.2.4)), one forward moving (related to I_{on}^{+}, eventually contributing to the total I_{on}) and one backward moving (i.e the $K_B T$-layer backscattered carriers, related to I_{on}^{-}). These two streams are assumed to have equilibrium Maxwellian distributions and have the average velocities v^{+} and v^{-}, both equal to the non-degenerate, unidirectional thermal velocity v_T.

   Natori first discussed the validity of this in [213]. More recent full band self-consistent Monte-Carlo simulations [214][188] show that while v^{+} was approximately v_T, v^{-} was only $\simeq 0.7 v_T$. If this is not accounted for, r_{sat} is overestimated by $\sim 10\%$ (and the I_{on} by $\sim 7{-}8\%$). In addition, the assumption that the positive and negative velocity carriers have a Maxwellian distribution has been shown to be inaccurate, esp. under high V_D conditions where strong non-equilibrium transport takes place.

2. Carriers injected into the channel from the source occupy states at the top of the barrier according to the Fermi level of the source. The $K_B T$-layer model assumes that the charge at the top of the barrier is $C_G(V_G - V_{TH})$.

3. Under low bias conditions, r is given by $r_{lin} = L/(L+\lambda)$, where λ is the momentum relaxation length related to the low field long channel mobility and L is the channel length.

   The expression for r_{lin} actually has a sound fundamental basis [215] in mesoscopic physics. It was also shown in [178] that this relation can be actually derived from the Drift-Diffusion formalism. The authors specify that the assumptions of the DD approach are valid for low V_D conditions even in a nanoscale device. There is some debate about this however, but for the most part, assuming that the relaxation time approximation (RTA) is valid is not off the mark. Thus usage of of a relaxation length λ related to μ_0 and v_T should be right to the first order.

4. Under high drain bias conditions, r is given by $r_{sat} = l/(l + \lambda)$, where l is the length of the critical region where the backscattering happens. It was assumed in [172] that this length was the width of the $K_B T$-layer l_{KT}.

   In 1979, Price observed [206] from some Monte Carlo simulations that carriers surpassing a potential barrier and travelling down the same potential were not likely to return to the injection point at the top of the barrier. Prior to Price's work, Bethe [216] showed that a near ballistic current occur in a forward biased metal-semiconductor junction, if the a potential drop equal to the $K_B T/q$ is seen before carriers can scatter. Thermionic emission is applicable in this case, due to the much shorter critical distance compared to the width of the barrier.

   Lundstrom reasoned that the critical layer for backscattering in a MOSFET is also roughly the distance over which the channel potential drops by $K_B T/q$, which is a tiny fraction of the channel length. Thus he extended the expression for backscattering in a field-free slab ($r = L/(L + \lambda)$ to a MOSFET (under a high drain bias V_D) as $r_{sat} = l/(l + \lambda)$ where l is the width of the $K_B T$ layer. Monte Carlo simulations in [175] show that actual critical layer is not quite the distance needed to drop $K_B T/q$ of potential, but is slightly higher. Lundstrom also later confirmed this result [217], noting that it is probably not possible to accurately estimate how much the width is higher without advanced simulations.

   In additon to the above points, it was also argued in [172] that because the critical backscattering occurs in a region where the carriers have gained little energy from the channel field, it was appropriate to use the same mean-free-path λ as used in the linear case above. The validity of this assumption has been questioned by Fischetti (see appendix A and B in [191]). In the same work, Fischetti also discusses source starvation and long range Coulomb effects and questions if it will be ever possible to attain ballistic transport proper in a MOSFET even if we scale down to sub-10 nm.

5. Scattering events deep inside the channel does not affect r, and thus I_{on}.
   It was argued that as a carrier moves deeper into the channel, its energy increases which raises its probability of scattering by phonon emission. When phonon emission happens beyond the KT-layer, the carrier energy is lowered and makes it less likely to return to the source. Even in the case of elastic scattering deep inside the channel, a carrier might undergo additional scattering events on its journey to the source. A sufficient number of these events will effectively thermalize the carrier energy and make it impossible to surmount the potential barrier. In addition, the interface roughness scattering is strong near the beginning of the channel and reduces as the carrier moves closer to the source – which means that the carriers far away from the source have a lesser chance to be elastically backscattered by this mechanism. Thus it was concluded that channel scattering does not change r as defined earlier, and thereby does not have an impact on I_{on}. However, in our opinion, this conclusion is inaccurate.

   Inelastic scattering inside the channel actually affects the height of the potential barrier between the source and channel through the coupling via Poisson's

equation. The potential barrier at the virtual source is not 'frozen' at its equilibrium, zero V_D value–it is affected by scattering in the channel [218]. The balance between forward-moving (v^+) and backward-going carriers (v^-) is lost once a current I_{on} starts flowing in the device. The surface potential at the virtual source will attempt to continuosly readjust, but it will never be able to come to its equilibrium value. Instead, a new equilibrium will be attained depending on the gate and drain voltages. The result of this new equilibrium is an increase in the barrier height (see, for instance Fig.(10) in [172] and Fig.(9) in [175]), which will result in a different value for r compared to the frozen field case. This feedback between barrier height and scattering is not taken into account in the $K_B T$-layer theory.

We strongly stress at this point that this effect due to channel scattering cannot be ignored today's MOSFETs. In the presence of strain, which affects material properties in a fundamental way, this feedback mechanism needs careful consideration, Natori has come up with a new theory explaining high field transport in semiconductors [192][193] , which delves into the origin of this feedback/coupling between inelastic scattering and backscattering. In later sections of this chapter, we show how the presence of this feedback is crucial in understanding and quantifying the nature of quasi-ballistic transport in a device.

7.2.7 Revisiting the Gildenblat Model

In 2002, Gildenblat presented [219] a slightly different approach to modeling the ballistic efficiency. Starting from the fundamental defintion of the scattering matrix (Section 7.2.3), Gildenblat developed a closed form analytical expression for the overall transmission coefficient for a device having an arbitrary potential profile. In the same paper, he also showed the LM model expression for r_{sat} can be recovered from his expression very easily. We refer to this transmission model as the Gildenblat Model (GM), upon which our model is based. For completeness, we outline the key points in Gildenblat's model in this section.

The starting point for GM are the equations presented in section 7.2.3. To arrive at an analytical expression for the transmission coefficient t, Gildenblat assumes that the ratio η is independent of position (while both μ^* and D^* are position and field

dependent). Starting from Eq. 7–15 and using

$$E_x = -\frac{d\phi}{dx} \tag{7–33a}$$

$$\frac{dn^*}{dx} = \frac{dn^*}{d\phi}\frac{d\phi}{dx} \tag{7–33b}$$

We have

$$\frac{J}{D^*} = -\left(-\eta n^* + \frac{dn^*}{d\phi}\right)\frac{d\phi}{dx} \tag{7–34}$$

Now consider the quantity $n^* e^{-\eta\phi}$, where η independent of ϕ. Differentiating this w.r.t ϕ, we have,

$$\frac{d}{d\phi}(n^* e^{-\eta\phi}) = e^{-\eta\phi}\frac{dn^*}{d\phi} - \eta e^{-\eta\phi} n^* \tag{7–35}$$

or

$$e^{\eta\phi}\frac{d}{d\phi}(n^* e^{-\eta\phi}) = \frac{dn^*}{d\phi} - \eta n^* \tag{7–36}$$

Using Eq.(7–36) in Eq. (7–34),

$$\frac{J}{D^*} = -e^{\eta\phi}\frac{d}{dx}(n^* e^{-\eta\phi}) \tag{7–37}$$

Rearranging, we get

$$\frac{d}{dx}(n^* e^{-\eta\phi}) = -e^{-\eta\phi}\frac{J}{D^*} \tag{7–38}$$

Noting that net flux J is a constant, and integrating both sides after adding appropriate limits,

$$\int_{n^*(0)e^{-\eta\phi(0)}}^{n^*(x)e^{-\eta\phi(x)}} d(n^* e^{-\eta\phi}) = -J\int_0^x e^{-\eta\phi}\frac{1}{D^*}dx \tag{7–39}$$

The result of the integration is

$$n^*(x)e^{-\eta\phi(x)} - n^*(0)e^{-\eta\phi(0)} = -J\int_0^x e^{-\eta\phi}\frac{dx}{D^*} \tag{7–40}$$

Here $\phi = \phi(x)$ denotes electrostatic potential

Defining $\psi(x) = \phi_(x) - \phi(0)$, it can be shown that

$$n^*(x) = n^*(0)e^{\eta\psi(x)} - Je^{\psi(x)} \int_0^x e^{-\eta\psi(x)} \frac{dx}{D^*} dx \tag{7–41}$$

Eq.(7–41) is a important result of [219]. Once $n^*(x)$ is known, the forward and the reverse fluxes can be easily estimated from Eq. (7–10) as

$$\begin{aligned} 2J^+(x) &= cn^*(0)e^{\eta\phi} + J\left(1 - ce^{\eta\psi(x)} \int_0^x e^{-\eta\psi(x)} \frac{dx}{D^*}\right) \\ 2J^-(x) &= cn^*(0)e^{\eta\phi} - J\left(1 + ce^{\eta\psi(x)} \int_0^x e^{-\eta\psi(x)} \frac{dx}{D^*}\right) \end{aligned}$$

It is well known from the flux method that the reflection and transmission coefficients of the system are given by

$$r = \frac{J^-(0)}{J^+(0)} \tag{7–43a}$$

$$t = 1 - r = \frac{J^+(L)}{J^+(0)} \tag{7–43b}$$

By defining an "effective velocity" ν_e given by

$$\frac{1}{\nu_e} = \int_0^L e^{-\eta\psi(x)} \frac{dx}{D^*} \tag{7–44}$$

and using the explicit equations for $J^+(x)$ and $J^-(x)$ shown above, it is easy to show that

$$J^-(0) = cn^*(0) - J \tag{7–45a}$$

$$J^+(0) = cn^*(0) + J \tag{7–45b}$$

$$\begin{aligned} J^+(L) &= cn^*(0)e^{\eta\psi(L)} \\ &+ J\left(1 - ce^{\eta\psi(L)} \int_0^L e^{-\eta\psi(x)} \frac{dx}{D^*}\right) \\ &= cn^*(0)e^{\eta\Delta\psi} + J\left(1 - \frac{c}{\nu_e} e^{\eta\Delta\psi}\right) \end{aligned} \tag{7–45c}$$

($\Delta\psi = \psi_L - \psi_0 = \psi(L)$ i.e. the surface potential difference between the drain and source ends for a particular voltage on the gate and drain terminals.)

Using the relation $r + t = 1$, we can show that

$$cn^*(0) = J\left(e^{-\eta\Delta\psi} + \frac{c}{\nu_e}\right) \tag{7–46}$$

Using the above expression, t can be expressed as

$$t = \frac{2}{1 + \frac{c}{\nu_e} + e^{-\eta\Delta\psi}} \tag{7–47}$$

Thus if ν_e as defined by Eq.(7–44) is known for a system, the transmission coefficient can be estimated using Eq.(7–47).

7.2.8 Revisiting the Original Shockley Work

Let us discuss the physics behind one-flux theory assumption in Eq. (7–12). In essence, this is a low field, near-equilibrium assumption, which is based on a simple linear dependence of the backscattering coefficients on the applied electric field. McKelvey [171] assumed the below expressions for the coefficients,

$$r^{\pm} = r_0(1 \pm \epsilon) \tag{7–48a}$$

$$r^{+} + r^{-} = 2r_0 \tag{7–48b}$$

$$r^{+} - r^{-} = (2r_0)\epsilon \tag{7–48c}$$

It is clearly stated in the initial McKelvey papers [169][171] that the method is applicable for “small applied electric fields” only. ϵ was shown to be a quantity linearly dependent on the electric field as

$$\epsilon = \frac{\mu_0}{c} E_x \tag{7–49}$$

(μ_0 is the low field mobility defined as $\mu_0 = e\tau_t/m$)

McKelvey quantified “small electric fields” by $\epsilon \ll 1$. Theoretically, this means that the carrier distribution function is not far away from equilibrium and the average carrier

velocity is small compared to c (i.e, velocity saturation, an high field effect, has not happened yet).

Thus, the assumption that $r^+ - r^- = \eta E_x$, with η being a constant – independent of position and independent of potential – Gildenblat's final expression for the transmission coefficient is essentially no different than what could be obtained from an Drift-Diffusion perspective. We reiterate that the Gildenblat result is simply a Lorentz solution of the Boltzmann Transport Equation. That is, the solution obtained by this method is simply the equilibrium distribution function upon which a drift velocity (arising from the presence of a small electric field) has been superposed. By its nature, this solution cannot describe quasi-ballistic transport with any more accuracy than a traditional DD solution. The Lundstrom quasi-ballistic transport model also suffers from a similar problem that shows up a different way, because of the same underlying assumption about the backscattering differences in the forward and reverse directions. Other authors have also discussed the validity of Lundstrom's assumptions from a different perspective - Vaidyanathan and Pulfrey present an good discussion of the issues in [190] and Fischetti [214] presents very comprehensive Monte Carlo simulation results of a nanoscale transistor, showing the off-equilibrium nature of the carrier distribution function even in the linear range, for voltages as small as 0.1V. Note that the M-flux theory developed by Lundstrom's group does not have these problem; the trade-off is model complexity.

7.3 UF Compact Model: Theory

To overcome the issues presented in the previous section, we have extended the Gildenblat approach for transmission by including high field transport effects and confinement effects (as would occur in the inversion layer of a MOSFET). One of the goals of our work was to retain analytical simplicity. We believe that including the confinement effects into the frame work of the one-flux theory is a indirect way of incorporating energy resolved fluxes for carrier transport. In this section, we first explain

the addition of high field effects, and then we show the addition of confinement effects to the resulting model. We also present a new interpretation of Natori's high field transport model [192][193], proving the equivalence of our approach and Natori's approach. With the help of this equivalence, we then explain the physical significance of the modified effective velocity term used in our model (in GM, it was simply used as a mathematical simplification). Finally we touch upon using our model for a clearer understand strain effects on quasi-ballistic transport from a new perspective.

7.3.1 Including High Field Effects

Our method for including high field effects stems from a suggestion put forward by Pulver and McKelvey in 1966 [220] in their paper discussing the application of the flux method for systems with nonconstant electric fields. In their work, they propose an empirical expression for the variation of $r^{\pm}$ with ϵ (and thereby E_x) that is applicable even at high fields. While not rigourously justified in a fundamental way, their expression reproduces the correct behavior of $r^{\pm}$ at all limits. The expression suggested is

$$r^{\pm} = r_0 e^{\pm\epsilon} \tag{7–50}$$

Under increasing electric fields, $\epsilon = \mu_0 E / c_x$ increases in value (and will eventually saturate when carrier velocity saturates) and eventually the condition $\epsilon \ll 1$ becomes invalid. In this limit, Pulver and McKelvey argue that, for a thin slab of thickness dx, $r^+ dx \rightarrow 0$ and $r^- dx \rightarrow 1$. They note that in extending these results to the limit $dx \rightarrow 0$, and thus, $r^+ \rightarrow 1$ and $r^- dx \rightarrow \infty$. it is also noted that the condition that $r^+(E) = -r^-(-E)$ should apply from symmetry considerations. The exponential relation suggested above is the simplest mathematical expression that satisfies both these conditions simultaneously.

With this form for the backscattering coefficients, the mobility and the diffusion expression become

$$\mu^*(x, E_x) = \left(\frac{r^+ - r^-}{r^+ + r^-}\right)\frac{c}{E_x} \tag{7–51}$$

$$= \frac{c}{E_x}\frac{e^{\epsilon} - e^{-\epsilon}}{e^{\epsilon} + e^{-\epsilon}} = \frac{c}{E_x}\tanh(\epsilon) \tag{7–52}$$

$$\tag{7–53}$$

$$D^*(x, E_x) = \frac{c}{r^+ + r^-} \tag{7–54}$$

$$= \frac{c}{r_0}\frac{1}{e^{\epsilon} + e^{-\epsilon}} = \frac{c}{2r_0}\frac{1}{\cosh(\epsilon)} \tag{7–55}$$

The expression for position/field dependent Einstein relation is,

$$\frac{D^*}{\mu^*} = \frac{E_x}{2r_0}\frac{1}{\sinh(\epsilon)} \tag{7–56}$$

$$\tag{7–57}$$

At the low field limit, we can verify that the above expression takes it default value –

$$\lim_{E_x \to 0}\frac{D^*}{\mu^*} = \frac{E_x}{2r_0}\frac{1}{\epsilon} \tag{7–58}$$

$$= \frac{c}{2r_0}\frac{1}{\mu_0} = \frac{D_0}{\mu_0} = \frac{k_B T}{q} \tag{7–59}$$

The expression for D^*/μ^* shown above depends on the electric field, which is the derivative of the electrostatic potential, but not on the electrostatic potential itself. This means that Gildenblat's mathematical approach is still applicable – specifically, Eq. 7–35 still works. This means that we can derive an expression for $n^*(x)$ using the same technique was as outlined in the last section after a (careful) integration involving a position dependent η. However, the "effective velocity" ν_e needs to be redefined to account for the position dependent η term. Once this is done, we can arrive at an expression for the transmission coefficient that reduces to the Gildenblat result in the low field limit. This procedure is outlined in the next section.

We start from Eq. 7–37 and use the term $\beta = \mu^*/D^*$ instead of η for clarity. We reiterate that the position and field dependence of β cannot be ignored.

$$\frac{J}{D^*} = -e^{\beta\phi}\frac{d}{dx}(n^* e^{-\beta\phi}) \tag{7–60}$$

Rearranging, we get

$$\frac{d}{dx}(n^* e^{-\beta\phi}) = -e^{-\beta\phi}\frac{J}{D^*} \tag{7–61}$$

or

$$\int d(n^* e^{-\beta\phi}) = \int -e^{-\beta\phi}\frac{J}{D^*}dx \tag{7–62}$$

Since β depends on position, we account for this in the limits of integration on the right hand side (β_0 and β_x, which are values of β when $x = 0$ and $x = x$, respectively).

Noting that net flux J is a constant and adding limits,

$$\int_{n^*(0)e^{-\beta_0\phi(0)}}^{n^*(x)e^{-\beta_x\phi(x)}} d(n^* e^{-\beta\phi}) = -J\int_0^x e^{-\beta\phi}\frac{1}{D^*}dx \tag{7–63}$$

After integration, we have,

$$n^*(x) = e^{\beta_x\phi(x)}\left(n^*(0)e^{-\beta_0\phi(0)} - J\int_0^x e^{-\beta\phi}\frac{dx}{D^*}dx\right) \tag{7–64}$$

Eq.(7–64) is the analog of Gildenblat's Eq.(7–41). We proceed in the same way as before to estimate $J^{\pm} = cn^*(x) \pm J$.

$$\begin{aligned} 2J^+(x) &= cn^*(0)e^{\beta_x\phi(x)}e^{\beta_0\phi(0)} \\ &+ J\left(1 - ce^{\beta_x\phi(x)}\int_0^x e^{-\beta\phi(x)}\frac{dx}{D^*}\right) \\ 2J^-(x) &= cn^*(0)e^{\beta_x\phi(x)}e^{\beta_0\phi(0)} \\ &- J\left(1 + ce^{\beta_x\phi(x)}\int_0^x e^{-\beta\phi(x)}\frac{dx}{D^*}\right) \end{aligned}$$

Redefining the "effective velocity" ν_e as below,

$$\frac{1}{\nu_e} = \int_0^L e^{-(\beta_x\phi(x)-\beta_0\phi(0))}\frac{dx}{D^*} \tag{7–66}$$

and following the same procedure as in the previous section, we can derive an expression for the transmission coefficient t as

$$t = \frac{2}{1 + \frac{c}{\nu_e} + e^{-(\beta_L \phi(L) - \beta_0 \phi(0))}} \tag{7–67}$$

If β is independent of postion/field (i.e. a constant), Gildenblat's Eq.(7–47) is recovered.

7.3.2 Including Confinement Effects

Unlike bulk devices, MOSFETs with 2D and 1D inversion layers operate under the effect of gate field induced quantum confinement [221]. The confinement of carriers lead to reduction of carrier degree of freedom. This in return results in confinement induced valley/band splitting, leading to repopulation of carriers between the bands. Additionally, the splitting of the bands modifies the joint density of states for scattering matrix elements [222]. For example, for Si electron inversion layers, the splitting of Δ_2 and Δ_4 valleys lead to reduction in f-type optical phonon scattering rates. Moving from a 3D to 2D bandstructure reduces the effective density states which also contributes to additional scattering rate reduction.

The gate-to-channel potential difference varies along the channel based on the conduction band profile. This leads to a position dependent level of quantum confinement for the carriers. At the source end, where the gate-to-channel potential difference is highest, the confinement induced bandsplitting is at its largest value. Hence, the carrier repopulation and the resultant conductivity mass modification (compared to bulk) is more pronounced compared to drain end. For instance, in (100)-oriented Si, the effective conductivity mass for electrons at the source end is lower compared to the drain end.

With a position dependent conductivity mass, we can define a position-dependent thermal velocity,

$$c(x) = \sqrt{\frac{2k_B T}{\pi m(x)^*}} \tag{7–68}$$

Adding the position dependence of thermal velocity to t_{UF}, we get

$$t_{UF} = \frac{2}{1 + \frac{c(0)}{c(L)}\left(\frac{c(L)}{\nu_e} + e^{-(\beta_L \phi(L) - \beta_0 \phi(0))}\right)} \tag{7–69}$$

where $c(0)$ and $c(L)$ represent the average thermal velocity at the source and drain end of the channel.

With the proper inclusion of high-field effects and confinement effects Eq. (7–69) becomes applicable for describing carrier transport for the entire range of device operation (low and high fields) and for all surface channel/orientations.

7.3.3 Revisiting the Natori's High Field Transport Model

In 2009, Natori presented a new approach [192][193] towards understanding high field transport in bulk semiconductors. In this approach, an analytical solution of a pseudo-one-dimensional BTE with a constant electric field is presented, after transforming it into a pair of carrier flux equations. Neither the relaxation time approximation nor the perturbation expansion is used in solving the equations. Explicit expressions for average carrier velocity and density as functions of field and position are presented.

The main concepts in Natori's theory are schematically presented in Figure 7-4. Natori considers a simple resistor in his papers, with a linear potential profile in the device "channel". The carriers entering the channel from the source has a distribution of energies, and an average value of $K_B T/q$ is assumed. In the region of the channel up to distance x_0 from the source, the carriers do not have enough energy to emit an optical phonon. In other words, the only scattering events in this region are elastic in nature (surface roughness, ionic impurity and acoustic phonon scattering). This region is therefore named as the Initial Elastic Zone (IEZ). In traversing the length of the IEZ,

the carriers gain energy from the field and have enough energy for optical phonon (OP) emission after crossing this zone.

Beyond the point $x = x_0$, OP emission happens in a very short time and carriers are allowed to relax to the lower energy level $\varepsilon - \varepsilon^*$. This energy level $\varepsilon - \varepsilon^*$ is known as the Ist Relaxed Level, and the carriers in the first relaxed level are only allowed to populate $x \geq x_0$. The carriers continue to gain energy from the field once again, till they reach the point $(x_0 + x_1)$ inside the channel where the carrier energy is equal to the OP energy. The region $x_0 \leq x \leq (x_0 + x_1)$ is known as the Ist Relaxed Elastic Zone.

It is easy to see that the entire channel can be represented as a cascade of these relaxed elastic zones with OP emission happening at each zone boundary. The total transmission through the device is therefore a product of the transmission coefficients through each elastic zone.

$$T = T_0 \cdot T_1 \cdot T_2 \cdots T_N \tag{7–70}$$

However, there is one difference between T_0 and the rest of the transmission coefficients $T_1 \cdots T_N$. Before x_0, energy relaxation can happen only through elastic scattering processes (acoustic phonon scattering process is considered to be elastic). Beyond this this point, carrier energy relaxation happens through both elastic and inelastic scattering mechanisms, and transport phenomena becomes conceptually identical in all the zones beyond IEZ. Thus, $T_1 \cdots T_N$ can be combined into one factor, and a single equivalent Optical Phonon Emission Zone can be visualized.

Right near the zone bounday x_0, after the inelastic scattering with OP, there is a chance for the carriers to be elastically backscattered into the IEZ. The probability for re-entry into the IEZ reduces as the carrier moves away from x_0 eventhough it can undergo an elastic scattering event at any further time. The total number of backscattered carriers into the IEZ depends on the amount of elastic backscattering and the energy relaxation via OP emission beyond $x \geq x_0$. Natori describes this through

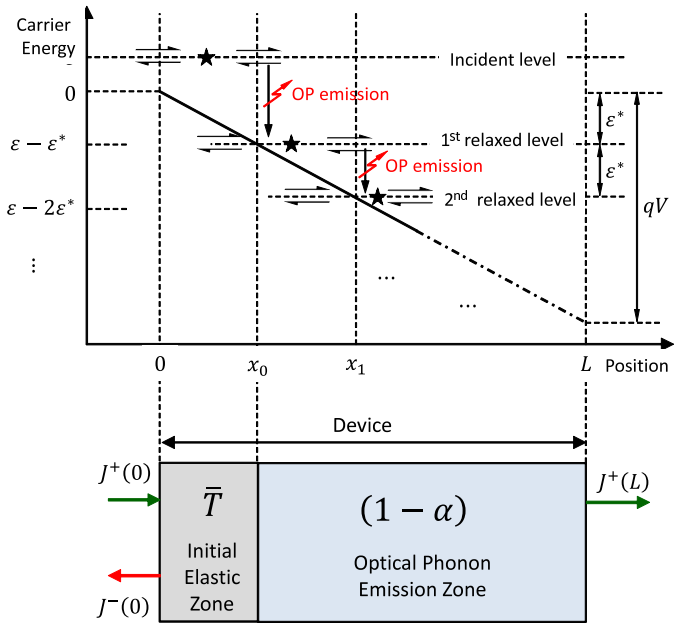


Figure 7-4. Natori's high field transport model showing the conceptual "Initial Elastic Zone" and "Optical Phonon Emission Zone" regions in a device carrying current.

a parameter α which represents the elastic-backscattering/energy relaxation trade-off. The presence of this tradeoff is very significant in understanding quasi-ballistic transport both with and without strain involved.

We would like to point out that the LM picture is slightly different in this regard, as the backscattering parameter r in that model represents the fraction of carriers getting back into the *source*, while Natori's α parameter accounts for backscattering from all of OPEZ into the *first elastic zone*. This is an subtle difference that has strong implications in the understanding of either model.

For a given energy of the incoming carriers, the IEZ width x_0 (and therefore the OPEZ width $L - x_0$) will depend on the applied voltage on the drain. Since the carriers will have a distribution of energies at the source, each energy will have it's own zone definitions for a given drain voltage as well. It is possible to conceptually think of a energy averaged IEZ and OPEZ for a given bias condition. High field transport in the

device is analyzed using a reflection-transmission method with these zones, similar to the one-flux approach.

Using very rigorous analysis and calculations, Natori showed that the current-field and the velocity-field relationships in this device were closely related to the flux transmission thorough the IEZ. If $\bar{T}$ and $(1-\alpha)$ represent the zone transmissions of the IEZ and OPEZ respectively, one would expect to see the overall transmission of the entire device to be the product of individual transmissions, i.e, $\bar{T}(1-\alpha)$ from the one-flux theory. However, Natori's analysis shows that the overall transmission is of the form shown in Eq. (7–71).

$$T_N = \frac{\bar{T}(1-\alpha)}{1-(1-\bar{T})\alpha} \tag{7–71}$$

This was attributed to the inherent feedback present between carrier transport and electrostatics through the Poisson's equation in the device. Note that the transmission will be a strong function of the resultant field in the channel. Using this transmission expression, Natori explains current saturation that happens under field conditions. Traditionally, current is thought to be the product of total carrier charge and the carrier velocity. Natori introduces the transmission into the device current equation as

$$I \propto qn_0 <v> \frac{\bar{T}(1-\alpha)}{1-(1-\bar{T})\alpha} \tag{7–72}$$

Using his transmission model, Natori points out that current through the device will saturate when the transmission $\bar{T}$ becomes equal to unity, which happens when zone width $x_0 \rightarrow 0$ at high field conditions. When this happens, the total transmission T_N saturates to the value $(1-\alpha)$. The level at which current saturates, therefore, depends on the transmission $(1-\alpha)$ of the OPEZ (which is the entire length of the channel as $x_0 \rightarrow 0$). We will explain what velocity is to be used in the current expression $<v>$ in later paragraphs.

7.3.4 Understanding Feedback in Natori's High Field Model

We first show that it is possible to arrive at Natori's expression for total transmission in a very simple and elegant manner using the one-flux theory, with the addition of carrier exchange between the zones. For simplicity, we use the notation $\bar{R} = 1 - \bar{T}$ from here on. Figure 7-5 shows a pictorial representation of our approach. The red arrows represent the reflected fluxes and the green arrows represent the transmitted fluxes from/to either zone in the device.

Consider an incident flux f_0 entering the channel; a part of this flux $(\bar{T} f_0)$ is transmitted through the IEZ, and part of it is reflected back into the source, namely $\bar{R} f_0$. Of the flux entering the OPEZ, the portion $(1 - \alpha)\bar{T} f_0$ is transmitted to the drain, and the fraction $\alpha \bar{T} f_0$ is reflected back into the IEZ. Note that these back reflections from the OPEZ includes the contribution from all the relaxed zones beyond the IEZ. A part of the flux reflected from the OPEZ to the IEZ $(1 - \bar{R}\alpha f_0)$ transmits back to the source, and a part of it $(\bar{R}\alpha \bar{T} f_0)$ is reflected back into the OPEZ. By considering carrier exchange through infinite reflection/transmission process that could happen between the two zones, we take into account the effect of inelastic scattering on the total transmission in a fundamental way.

To derive the expression for the total transmission, we need three terms: (1) incident flux $J^+(0)$ (2) total transmitted flux $J^+(L)$ (3) total reflected flux $J^-(0)$. The flux terms can be obtained by summing the components shown in Figure 7-5 as,

$$
\begin{aligned}
J^{+}(0) &= f_0 \\
\\
J^{-}(0) &= \bar{R} f_0 + (1-\bar{R})\alpha \bar{T} f_0 + (1-\bar{R})\bar{R}\alpha^2 \bar{T} f_0 \\
&\quad + (1-\bar{R})\bar{R}^2\alpha^3 \bar{T} f_0 + (1-\bar{R})\bar{R}^3\alpha^4 \bar{T} f_0 + \ldots \\
\\
J^{+}(L) &= (1-\alpha)\bar{T} f_0 + (1-\alpha)\bar{R}\alpha \bar{T} f_0 \\
&\quad + (1-\alpha)\bar{R}^2\alpha^2 \bar{T} f_0 + \ldots
\end{aligned}
\tag{7–73}
$$

Summing up the infinite series for $J^{+}(0)$ and $J^{-}(0)$, we have

$$
\begin{aligned}
J^{-}(0) &= \bar{R} f_0 + (1-\bar{R})\alpha \bar{T} f_0 \sum_{n=0}^{n=\infty} \bar{R}^n \alpha^n \\
&= \bar{R} f_0 + (1-\bar{R})\alpha \bar{T} f_0 \frac{1}{(1-\bar{R}\alpha)} \qquad (7–74) \\
J^{+}(L) &= (1-\alpha)\bar{T} f_0 \sum_{n=0}^{n=\infty} \bar{R}^n \alpha^n \\
&= (1-\alpha)\bar{T} f_0 \frac{1}{(1-\bar{R}\alpha)} \qquad (7–75)
\end{aligned}
$$

and finally,

$$
T_N = \frac{J^{+}(L)}{J^{+}(0)} = \frac{\bar{T}}{1 + \dfrac{\alpha}{(1-\alpha)}\bar{T}} \tag{7–76}
$$

which is identical to Natori's expression (Eq. (7–71)). A similar result is also seen in many other fields, for example in the treatment of optical transmission through multilayered structures [223], when interference effects are neglected and all multiple internal reflections are taken into account, similar to our case.

The form of the expression for the total transmission in Eq. (7–76) is similar to the transfer function of a negative feedback system commonly encountered in control theory [224]. This analogy is not coincidence, but rather, a result of the effect of the feedback

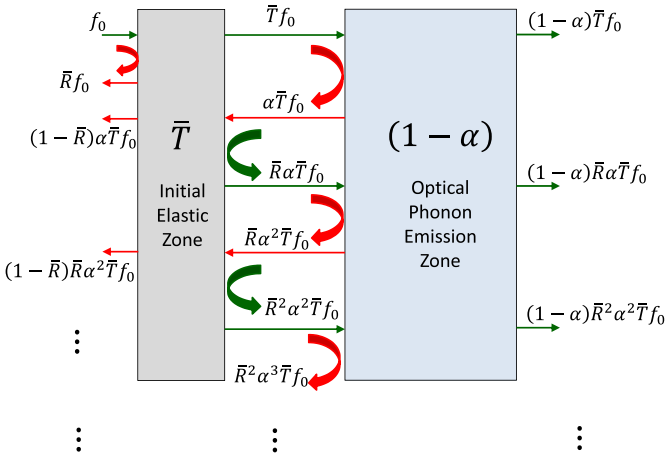


Figure 7-5. Our derivation of the overall transmission coefficient using an infinite series summation approach.

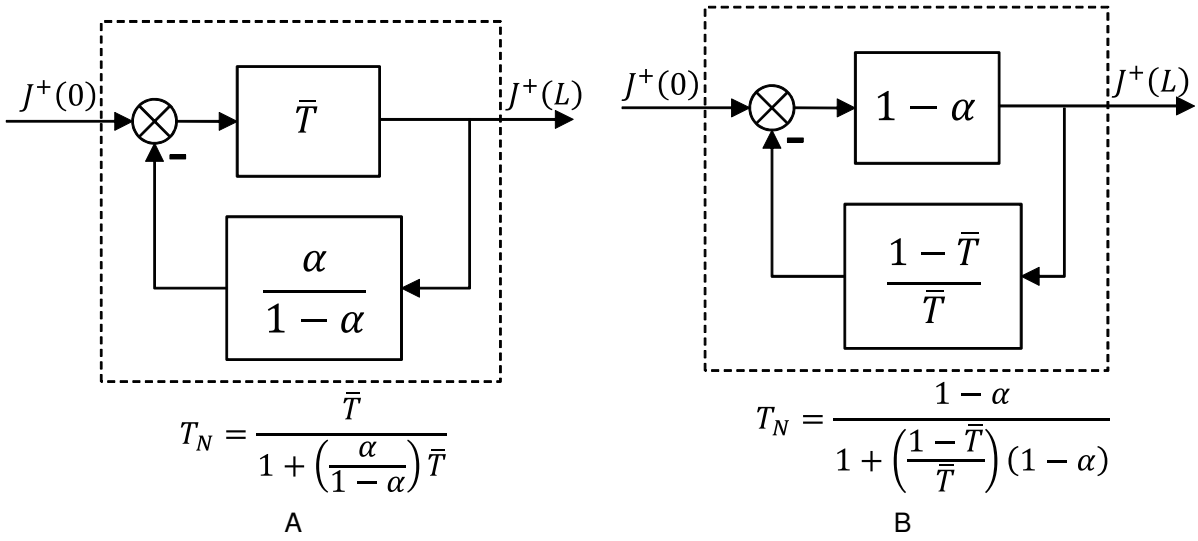


Figure 7-6. Two equivalent negative feedback system representation of the Natori Model with (a) FEZ transmission in the feed-forward path and (b) OPEZ transmission in the feed-forward path showing how the feedback from the OPEZ regulates the total current through the device.

between carrier transport and the potential inside the device. In [193], Natori explains how this feedback relates to the current control mechanism in the device, and ultimately is responsible for velocity and current saturation in the device.

In general, the transmission coefficients of the IEZ and OPEZ are inversely related to their respective widths. As the applied field increases, the IEZ width $x_0 \rightarrow 0$ and

the OPEZ width $\rightarrow L$, thereby making $\bar{T}$ and $(1-\alpha)$ proportional to the applied field. The exact nature of the field dependence will be related the potential profile inside the device. With this understanding, it is easy to see that the $I-V$ characteristics will be closely related to the carrier transmission through these zones. For a linear potential profile discussed by Natori, when the applied field is the range such that the transmission coefficient $\bar{T} \ll 1$, Eq. (7–72) results in the familiar Ohm's law (current linearly proportional to applied field).

When the applied field is increased such that $\bar{T} \rightarrow 1$, the current and the *average* carrier velocity both saturate. Natori's theory anticipates that the instantaneous velocity of the carrier is not constant from its transit from the source to drain. Instead, it periodically oscillates in each of the zones as the carrier is successively transmitted to the 1st, 2nd and higher relaxed levels. However, the mean carrier velocity across all the zones $< v >$ is uniform and remains proportional to the electric field up to 10^4 V/cm, beyond which it saturates. The carrier density, on the other hand, also tends to have periodic nature across the zones once a current starts flowing. To maintain current continuity, the $n(x)v(x)$ product should be constant at all positions inside the device. Since the position averaged carrier velocity $< v >$ is found to be a constant (like in Ryder's work [225]), the carrier concentration multiplying this velocity should be a constant too. Natori shows that the carrier concentration to be used in the current equation is the equilibrium carrier density n_0, which is set by the feedback process from OPEZ. The (negative) feedback acts to minimize the total electrostatic energy of the system via Poisson's equation.

To understand the feedback, we first remind that the OPEZ is a repetition of the same unit structure of width x_1 as shown in Figure 7-4. For the given field condition, the resultant current is set in the IEZ, the flux distribution in the 1st and other higher relaxed levels are determined in accordance with the carrier velocity in the respective zones so as to maintain current continuity. However, since a current is flowing, the

carrier distribution along the zones are not be consistent with the original constant field distribution.

The carrier charge excess (deficit) pushes up (pulls down) the potential profile of the original constant-field curve, which modifies the profile (i.e the “barrier height”) at the beginning of the channel. This results in a decrease (increase) of the carrier injection, and thus compensates for the excess (deficit) of the charge. The potential profile in the IEZ is modified by the feedback from the rest of the channel. The steady-state potential profile is obtained by the self-consistent solution of the coupled system of the BTE and the Poisson equation. This self-consistency supports (i.e., tries to bring the system back to) the constant field condition forced externally by the applied bias. The carrier distribution in the bulk acts in conjunction with the dopant charge distribution (if the device is doped) to minimize the total electrostatic energy through the feedback mechanism. The feedback results in setting the average carrier distribution $<n>= n_0$ and the current becomes proportional to $qn_0 <v> T_N$.

The conventional theory of velocity saturation is based on the balance-of-energy equation and predicts that the carrier velocity will saturate if the optical phonon scattering dominates the energy relaxation, which then causes the current to saturate. Natori’s model shows that the velocity is proportional to E if the transmission is small, even if the optical phonon emission is dominant. If $\bar{T} \rightarrow 1$, then mean carrier velocity through the entire device saturates and causes the current to saturate. Saturation would happen irrespective of the field value if somehow we were to make $\bar{T} \rightarrow 1$ happen. In other words, transmission saturation through the IEZ causes current saturation (to the level determined by product of qn_0, $<v>$ and $(1-\alpha)$). This is a fundamental difference between conventional models and Natori’s new model.

Two equivalent representations of a negative feedback system with a transfer function as Eq. (7–76) are shown in Figure 7-6. Without feedback, the system representation would be a simple cascade of $\bar{T}$ and $(1-\alpha)$. The stabilizing negative

feedback from the OPEZ acting through α helps to regulate the total current the device. While Figure 7-6A describes the actual physical picture, model Figure 7-6B helps understand the saturation of total transmission when the length of the first elastic zone becomes so small that the $\bar{\bar{T}}$ becomes unity.

Our method of using the one-flux theory including infinite carrier exchange between the IEZ and OPEZ (Figure 7-5 and Eq. (7–76)) to derive the Natori model transmission equations facilitate a better link to the Lundstrom or Gildenblat compact models. While Natori discussed a bulk transport case in his paper, our derivation of Eq. (7–76) is quite generic and there is nothing precluding its applications to studying quasi-ballistic transport in a nanoscale MOSFET. Since our updated transmission coefficient including high-field and confinement effects (Eq. (7–69)) is a surface potential based equation, we can now equate it to the Natori equation Eq. (7–76), noting the inherent link and feedback present between the Possion equation, BTE and the resultant surface potential.

Using the feedback representation of the Natori expression shown in Figure 7-6A, we have

$$t_{UF} = T_N = \frac{\bar{\bar{T}}}{1 + \frac{\alpha}{(1-\alpha)}\bar{\bar{T}}} \tag{7–77}$$

$$1 - \alpha = \left(1 + \frac{1}{t_{UF}} - \frac{1}{\bar{\bar{T}}}\right)^{-1} \tag{7–78}$$

If t_{UF} can be computed for a real device, then Eq. (7–78) gives us a handle to estimate α for high bias conditions when $\bar{\bar{T}}$ saturates. One possible scheme for evaluating t_{UF} would be the technique proposed by Gildenblat [226] involving the PSP compact model [227][228], combined with simulation results for $c(0)$ and $c(L)$ from a Schroedinger-Poisson solution.

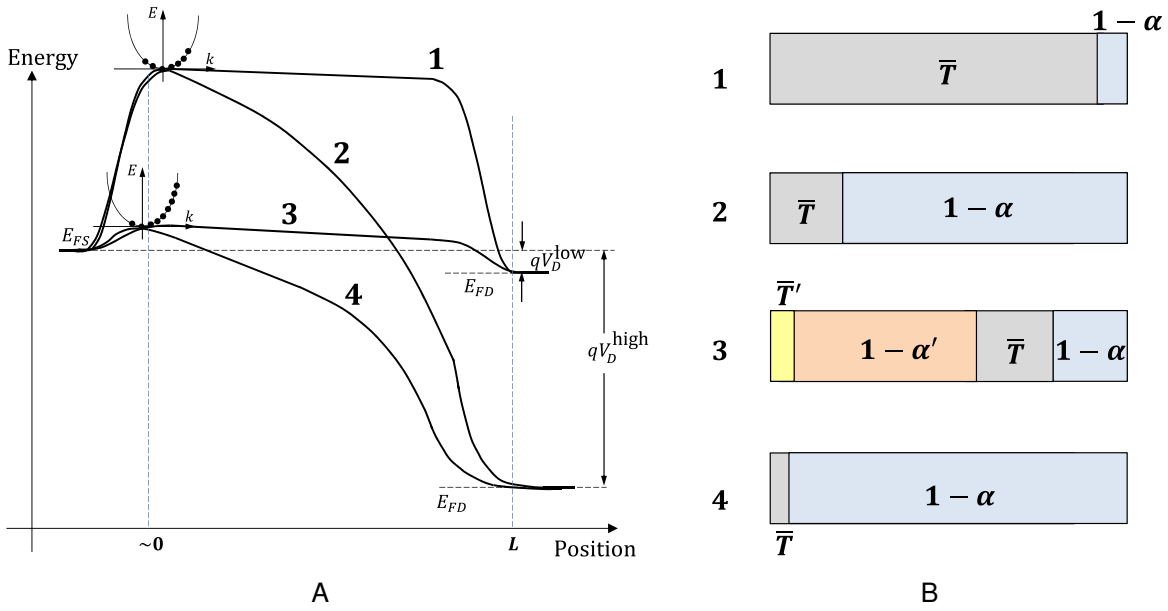


Figure 7-7. Evolution of (a) conduction band profile and (b) corresponding transmission characteristics of the channel are qualitatively shown for four possible device operating conditions: (1) low V_G, low V_D (2) low V_G, high V_D (3) high V_G, low V_D and (4) high V_G, high V_D

7.3.5 Physical Significance of Effective Velocity Term

In Gildenblat's original work [226], the effective velocity term (Eq. (7–44)) was simply a mathematical simplification used to simplify the transmission expression. However, our understanding of (1) the feedback mechanism and (2) the equivalence between the Natori expression and our updated expression for transmission provides a new physical understanding of the effective velocity term.

Since the redefined effective velocity is a definite integral over the channel length, we can split into two parts, with the splitting point being the FEZ width x_0 (for the given bias conditions) as,

$$\begin{aligned}\frac{1}{\nu_e} &= \int_0^L e^{-(\beta_x\phi(x)-\beta_0\phi(0))}\frac{dx}{D^*} \\ &= \int_0^{x_0} e^{-(\beta_x\phi(x)-\beta_0\phi(0))}\frac{dx}{D^*} \\ &\quad + \int_{x_0}^L e^{-(\beta_x\phi(x)-\beta_0\phi(0))}\frac{dx}{D^*}\end{aligned} \tag{7–79}$$

or,

$$\frac{1}{\nu_e} = \frac{1}{\nu_{e1}} + \frac{1}{\nu_{e2}} \tag{7–80}$$

The two equivalent transmission expressions we have are,

$$T_N = \left(\frac{1}{\bar{T}} + \frac{\alpha}{1-\alpha}\right)^{-1} \tag{7–81}$$

$$t_{UF} = \left(\frac{1}{2} + \frac{c(0)}{2c(L)}\left(\frac{c(L)}{\nu_e} + K\right)\right)^{-1} \tag{7–82}$$

where $K = e^{-(\beta_L\phi(L)-\beta_0\phi(0))}$.

Equating Eqs. (7–81) and (7–82), we get,

$$\frac{1}{\bar{T}} + \frac{\alpha}{1-\alpha} = \frac{c(0)}{2\nu_e} + \frac{K}{2}\frac{c(0)}{c(L)} + \frac{1}{2} \tag{7–83}$$

Expanding the velocity integral term, we have,

$$\frac{1}{\bar{T}} + \frac{\alpha}{1-\alpha} = \frac{c(0)}{2}\left(\frac{1}{\nu_{e1}} + \frac{1}{\nu_{e2}}\right) + \frac{K}{2}\frac{c(0)}{c(L)} + \frac{1}{2} \tag{7–84}$$

When the applied gate and drain voltages are high, the K term becomes much smaller than unity and can be neglected (Section IV of [219]). Since $x_0 \to 0$ under these voltage conditions, $\bar{T} \to 1$ (i.e., ν_{e1} term vanishes), making ν_e to be entirely composed of the ν_{e2} term, and . Therefore we have,

$$\frac{\alpha}{1-\alpha} = \frac{c(0)}{2}\left(\frac{1}{\nu_e}\right) - \frac{1}{2} \tag{7–85}$$

Solving for α (high bias), we get,

$$\alpha = \frac{1 - \frac{\nu_e}{c(0)}}{1 + \frac{\nu_e}{c(0)}} \tag{7–86}$$

Looking at the form of the ν_e term in Eq. (7–79), we can see that it represents a diffusion-velocity like quantity. In the absence of a drain bias, ν_e describes how efficiently the carriers would diffuse due to a given concentration gradient along the channel. The maximum diffusion velocity is obtained under the complete absence of scattering events, and is equal to the thermal velocity of the carriers $c(0)$. ν_e, therefore, can at the most be $c(0)$. When $\nu_e/c(0)$ becomes less than 1, it indicates the presence of the scattering events in the channel. Under an applied drain bias, the effectiveness of the superimposed drift response of the carriers depend on the diffusibility of the channel, i.e., the ease with which carrier transport happens. This is described by the $\nu_e/c(0)$ ratio. With this interpretation, we can readily see how scattering influences the overall device current. The impact of scattering events on the current is explained through the parameter α.

In a quasi-ballistic device, the saturation current is proportional to $(1-\alpha)$. From Eq. (7–86), when the $\nu_e = c(0)$, the elastic-backscattering/energy relaxation trade-off between the zones vanish making $\alpha = 0$. In other words, both parts of the channel have unity transmission, making the transport completely ballistic. Thus, we can now understand the physical significance of the effective velocity term. Using this concept, we present a new definition of ballistic efficiency in the device as,

$$B = \frac{<v>\,(1-\alpha)}{v_F} \tag{7–87}$$

where $<v>$ represents the position and energy averaged drift velocity,

$$<v> = \frac{\int_0^x \int_0^E n(x,E)\,v(x,E)\,dEdx}{\int_0^x \int_0^E n(x,E)\,dEdx} \tag{7–88}$$

and v_F represents the Fermi velocity at the beginning of the channel. When α becomes zero, carriers at the top of the barrier enter the channel with Fermi velocity v_F. When there are no scattering events along the channel, $<v>= v_F$, making $B = 1$. Note that current is still limited by number of available states at the beginning of the channel.

7.3.6 Expanding Natori's High Field Model for MOSFETs

While Natori's idea of dividing the channel into two zones with distinct transmission properties is quite generic, the application of this model for 2D or 1D inversion layers needs the inclusion of the effects of gate to channel potential. Figure 7-7A shows the evolution of the conduction band profile for various combinations of applied gate and drain biases. At the top of the barrier, the E-K diagram and occupation of electrons are illustrated. Carriers can attain kinetic energy due to drain bias while drifting in the channel, as well as from the degenerate occupation of the conduction band due to high gate bias. Accordingly the separation of the channel in terms of transmission characteristics are shown for all four cases in Figure 7-7B.

Case 1: Here the applied gate and drain voltages are low (weak inversion, linear regime). The potential profile in the channel is almost linear and the carrier statistics everywhere in the channel are non-degenerate. Hence, the kinetic energy of the carriers remains constant (near the thermal energy) making OP scattering dominant only at the drain entrance. The transmission picture is therefore simple, with the FEZ expanding to cover nearly the entire channel.

Case 2: When the drain voltage is increased keeping the gate voltage still low (weak inversion, saturation regime), the kinetic energy of the non-degenerate carriers in the inversion layer increases along the channel. This in return, increases the OP

scattering probability inside the channel, making the OPEZ cover a larger portion of the channel.

Case 3: Here we consider a large gate voltage and small drain voltage applied to the device (strong inversion, linear regime). The degenerate carriers at the top of the barrier have kinetic energy from the gate potential, and a fraction of these carriers will relax through OP scattering close to the beginning of the channel. The primed quantities ($\bar{T}'$ and $1-\alpha'$) describe the zone transmissions representing energy relaxation process of degenerate carriers only. When gate to channel potential difference reduces to the non-degenerate conditions, the carrier kinetic energy reduces to near thermal energies, and carrier transport from this point resembles case 1.

Case 4: With increase of drain bias (strong inversion, saturation regime) from case 3, the reduction of the kinetic energy of the degenerate carriers is compensated by the energy gained from high drain field. The OPEZ zones in case 3 merge (since transmission through the middle $\bar{T}$ zone saturates) and the channel in these conditions is almost entirely comprised of one OPEZ. This leads to a total transmission of $(1-\alpha)$ as predicted by Natori.

7.4 Qualitative Understanding of Strain Effects

Strain changes bond angles and lengths, which modifies the conduction and valence bands. These modifications include band warping and band splitting, resulting in modification of conductivity mass and scattering rates [229][230]. For electron inversion layers, the uniaxial stress effects [231] comes from: (1) conductivity mass decrease from the electrons repopulated from Δ_4 to Δ_2 valleys (2) OP scattering suppression due to the density of states decrease from Δ_4 to Δ_2 (3) conductivity mass decrease due to the band warping. For hole inversion layers, the stress effects [232] [233] arises from splitting and warping of heavy-hole and light-hole bands that results in (1) significantly reduced hole conductivity mass and (2) significantly reduced OP scattering.

There have been several experimental works [234][235] published that show the difference between linear current and saturation current enhancement with uniaxial stress under strong inversion conditions. It has been observed that for holes, the $I_{D(\mathrm{sat})}$ enhancement drops by nearly 50% compared to $I_{D(\mathrm{lin})}$, and while $I_{D(\mathrm{sat})}$ enhancement is lower for electrons, the discrepancy is not as much.

Looking at the transport characteristics shown in Figure 7-7B, the difference between carrier transmissions under linear and saturation conditions are immediately seen. Since the current is proportional to the product of average drift velocity and total transmission, the simultaneous impact of strain on each of these two parameters for linear/saturation cases need to considered. The velocity term is inversely proportional to the square root of the conductivity mass. The increase in the velocity term strictly depends on the strain induced reduction of effective conductivity mass. The strain dependence of the transmission term is somewhat more complex, since it is dependent on both both scattering and mass.

Under favorable stress conditions, inversion layer electrons experience reduced conductivity mass with minimal changes on scattering. But for holes, the favorable strain results in significant reduction of both carrier mass and scattering. From the strain perspective, the benefits of improved average carrier drift velocity is felt in both linear and saturation regimes. The differences in $I_{D(\mathrm{lin})}$ and $I_{D(\mathrm{sat})}$ enhancements therefore, must come from the differences in how strain affects the total transmission for the respective regimes,for both types of carriers.

Let us consider the linear regime first. For a high gate bias in the linear regime (Case 3 in previous section), the initial FEZ region becomes really small (i.e $\bar{T}' \rightarrow 1$). The channel then is comprised of three regions, an initial OPEZ characterized by $1-\alpha'$, the middle elastic region characterized by $\bar{T}$ and the final initial OPEZ characterized by $1-\alpha$. Carrier interchange occurs at both zone boundaries.

For holes, strain significantly reduces the OP scattering, which increases α and reduces total transmission. Since the mass reduction from strain induced bandwarping is significant, the increase in drift velocity is much bigger and compensates the strain induced transmission loss. Most of the enhancement in hole $I_{D(\mathrm{lin})}$, therefore, comes from the average drift velocity increase. For electrons, the effect of strain on scattering is small, and the corresponding transmission loss is also small. Most of the linear current enhancement for electrons comes from the average drift velocity increase as well. The amount of velocity increase, however, is smaller for holes compared to electrons.

In the saturation regime, the middle elastic zone transmission $\bar{T}$ saturates and the OPEZ zones merge together. The strain impact on velocity is still present for both type of carriers. OP scattering happens in a larger portion of the channel. For holes, the reduction in transmission from strain is much more compared to linear regime, resulting in a reduced enhancement in $I_{D(\mathrm{sat})}$. For electrons, the scattering reduction is small (compared to linear regime) and it's effect on current is not felt as much. Thus the electron $I_{D(\mathrm{sat})}$ also reduces, but not as much compared to hole $I_{D(\mathrm{sat})}$.

7.5 Summary

In this chapter, we reviewed the fundamentals of several significant models describing quasi-ballistic transport. The one-flux theory based Lundstrom Model and Gildenblat Model were examined in detail. By adding high field effects and confinement effects, leading to a new, surface potential based, analytical expression for the transmission coefficient in a MOSFET channel was established. This expression is applicable for all conditions of device operation and for all suface/channel orientations. The physical significance of the effective velocity was highlighted and it was shown to be an important figure of merit for describing the transport in a MOSFET channel. The essential physics of Natori's high field transport model was explained using a simple mathematical treatment and the model was extended to include the effect of gate confinement. By linking these two updated models, experimentally reported differences

between strain enhanced linear and saturation currents in short channel devices was qualitatively explained.

CHAPTER 8
OPTICAL CHARACTERIZATION OF ALGAN/GAN HEMT TRAP GENERATION

8.1 Motivation

After its first demonstration in 1994 [14], AlGaN/GaN HEMT devices gained a well deserved popularity for its high frequency, high power operation with high temperature stability, especially for high power microwave applications such as communications and radars. In 1996, the first reported microwave power operation of the AlGaN/GaN HEMT was 1.1W/mm power density at 2GHz operation frequency [236]. The continuous interest on high frequency, high power transistors boosted research efforts with a 30W/mm HEMT operating at 8GHz demonstrated in 2004 [17]. While the progress on the AlGaN/GaN HEMT performance improvement has been sustained, high power and high field operation have exacerbated reliability issues of GaN based HEMT devices which has limited the achievement of the complete potential of these devices [21].

As we have discussed in Chapter 3, the abundance of defects in the GaN material system can create trapping centers in the bulk GaN region, in the AlGaN barrier, at the AlGaN/GaN interface, at the Ni/AlGaN gate metal interface, and at the ungated region of the AlGaN surface [21–23]. In addition to the native defects formed during crystal growth and device fabrication, additional defects can be generated during device operation. One of the theories for degradation of AlGaN/GaN HEMT devices is the structural defect formation in the AlGaN barrier due to the addition of the inverse piezoelectric stress at high gate-to-drain electric fields to the large built-in epitaxial stress [20, 26]. The structural defect in the AlGaN layer caused by the increased stress is believed to create a leakage path from gate to channel. Another theory for the increased gate current is the metal diffusion into the AlGaN barrier at high gate fields, which is also known as “gate sinking” [21]. The reduced thickness in the AlGaN barrier where the gate “sinks” creates a leakage path by narrowing the AlGaN potential barrier. Further reliability and stability issues are current collapse, drain lag, and gate lag. Current collapse occurs

due to the trapping of hot electrons at high V_{DS} conditions in the slow deep bulk traps, in the AlGaN barrier or at the AlGaN surface [21–25]. These trapped electrons change the electrostatics of the channel and locally deplete the 2DEG. This in turn leads to a time dependent decrease in the drain current. Once the drain current is collapsed, the thermal emission of electrons from the trap sites is required for recovery of the current collapse. Hence, if the drain bias is pulsed from 0V to a high voltage and back to 0V, the drain current lags due to the trapping and detrapping processes related to current collapse (drain lag). Similarly, the electrons trapped on the AlGaN surface states creates a negative voltage and acts as a virtual gate on the device. When the gate voltage is pulsed from depletion to ON-condition, the drain current lags due to the emission time constant of the filled surface AlGaN states, i.e. the opening of the virtual gate near the drain side of the channel which results in gate lag.

Clearly, built-in traps, structural defect formation, and trap generation in AlGaN/GaN HEMTs are still important issues that need to be studied and solved. In addition to the built-in epitaxial stress ($\sim$ 3GPa for 28% Al concentration), the internal mechanical stress created during device operation due to the thermal expansion coefficient mismatch and by the inverse piezoelectric effect in the AlGaN barrier may add to the built-in stress and generate structural defects [20, 26]. It is also known that large local hydrostatic strain fields are induced in the lattice by point defects in GaN based materials [237–240]. Therefore, internal mechanical stress can alter the generation rate of point defects, such as Ga and N vacancies (V_{Ga} and V_N), interstitial Ga and N atoms in the lattice (Ga_i and N_i), substitutional O on N site (O_N), as well as the diffusion of the gate metal into the AlGaN layer. Hence, electrically active trap characteristics are expected to change over long periods of nominal high power operation due to a combination of electrical and internal mechanical stressing at elevated temperature. To decouple the contribution of mechanical stress in AlGaN/GaN HEMT degradation, a systematic study of applied external mechanical stress during device degradation

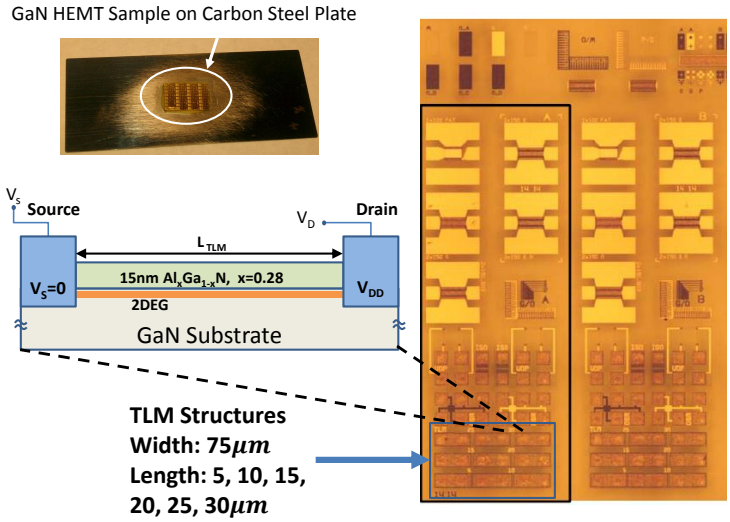


Figure 8-1. AlGaN/GaN HEMT on SiC wafer piece is attached to a carbon steel plate with epoxy (top left). The subreticles are shown on the right. The TLM module is located at the bottom of each subreticle. The cross sectional sketch of the AlGaN/GaN heterostructure, 2DEG and source/drain contacts are shown on the bottom left.

and trap characterization is required. Consequently, it is crucial to incorporate external mechanical stressing into established trap characterization methods to investigate the effect of mechanical stress on AlGaN/GaN trap generation. To achieve this goal, we developed a trap characterization method based on the photoionization spectroscopy technique, integrated into combined electrical and electrical and mechanical stressing experiments for controlled trap generation and characterization in one single device.

8.2 Overview of HEMT Device

In this study, AlGaN/GaN test wafers are used to investigate the mechanical stress effects on AlGaN/GaN HEMT trap generation. The test wafer layout consists of four subreticles, A to D. Each subreticle is populated with one FATFET, one center gate short channel HEMT, one off-center gate short channel HEMT, one 1μm long channel HEMT, one ungated HEMT site, isolation structures, van der Pauw test structures and transfer length module (TLM) structures. Among the subreticles, the length of the short channel HEMT varies as (A) 100nm, (B) 125nm, (C) 140nm and (D) 170nm.

In order to effectively photoionize the traps in the AlGaN/GaN layers without blocking the incident photons by the gate metal, we choose the TLM structures and shortest channel off-center gate HEMT devices to be included in our study. The TLM structures are located at the bottom of each subreticle in the current layout (Figure 8-1). The TLM structures are patterned in a 3-by-2 layout with a row separation of 50μm. The length of the TLM structures vary from 5μm to 30μm with increments of 5μm. All TLM devices have a fixed width of 75μm. For the short-channel off-center gate HEMTs, the gate-to-drain spacing is 3μm and the gate-to-source spacing is 1μm. Since the portion of the channel covered by the gate metal is minimal (0.1μm), these devices can also be effectively photoionized by the incident light. The epitaxial AlGaN/GaN heterostructure is grown uniformly throughout the whole wafer. Hence, both the TLM and HEMT sites are composed of a 15nm thick $Al_{0.28}Ga_{0.72}N$ barrier with a 2.25μm thick Fe-doped GaN buffer layer on top of an AlN nucleation layer. The bottom substrate is either a semi-insulating Si or SiC substrate. The source and drain Ti/Al/Ni/Au ohmic contacts are formed after a 20 seconds anneal at $850C^0$. All devices have a wide band gap SiN passivation layer on top, which is transparent to the wavelengths considered in this study.

The built-in 2DEG density for the SiC substrate samples is measured from Van der Pauw structures and found to be $1 \cdot 10^{13} cm^{-2}$. Additionally, we have also conducted capacitance-voltage measurements on the 125μm diameter circular CV-dots (Figure 8-2-a). By integrating the capacitance for a bias range of -6V (depletion) to 0V, we can obtain the density of the 2DEG (Figure 8-2-b). It is found that the SiC substrate devices have a higher 2DEG for a given gate bias compared to Si substrate devices due to a lower threshold voltage. The 2DEG density obtained from the CV measurements at 0V for the SiC devices is similar to the value extracted from the Hall measurements. This indicates a minimal effect of metal workfunction on the surface

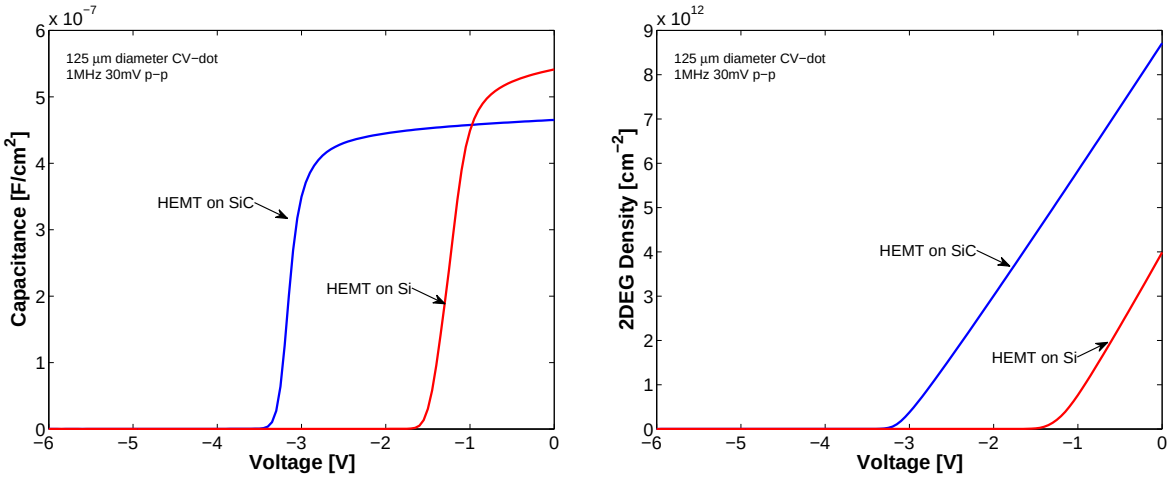


Figure 8-2. Capacitance-voltage measurement results (a) and the extracted 2DEG density (b) are shown for AlGaN/GaN HEMT on SiC and Si CV-dots. The threshold voltage difference between two samples create a significant 2DEG density at a given gate bias.

potenial at the AlGaN/GaN interface. Hence, the 2DEG density of ungated devices can be approximated by the 2DEG obtained from CV-dots at 0V gate bias.

8.3 Method

The AlGaN/GaN HEMT trap generation is investigated by a combination of controlled trap generation and photoionization spectroscopy trap characterization experiments. The controlled trap generation (degradation) in the device is performed by stressing the device while monitoring sudden increases in the gate current and degradation in the drain current. The energy distribution of areal trap density is obtained by measuring the drain current transients due to the additional photoionized electron emission from the trap states at different subbandgap wavelengths ($\geq$365nm). By restricting the photon energies below the GaN energy band gap (3.4eV), the optical excitation is limited to the subbandgap trap states and no new electron-hole pairs are generated.

As we have discussed in Section 3.2.6, photoionization spectroscopy previously has been employed for the investigation of slow bulk GaN traps and impurity sites via

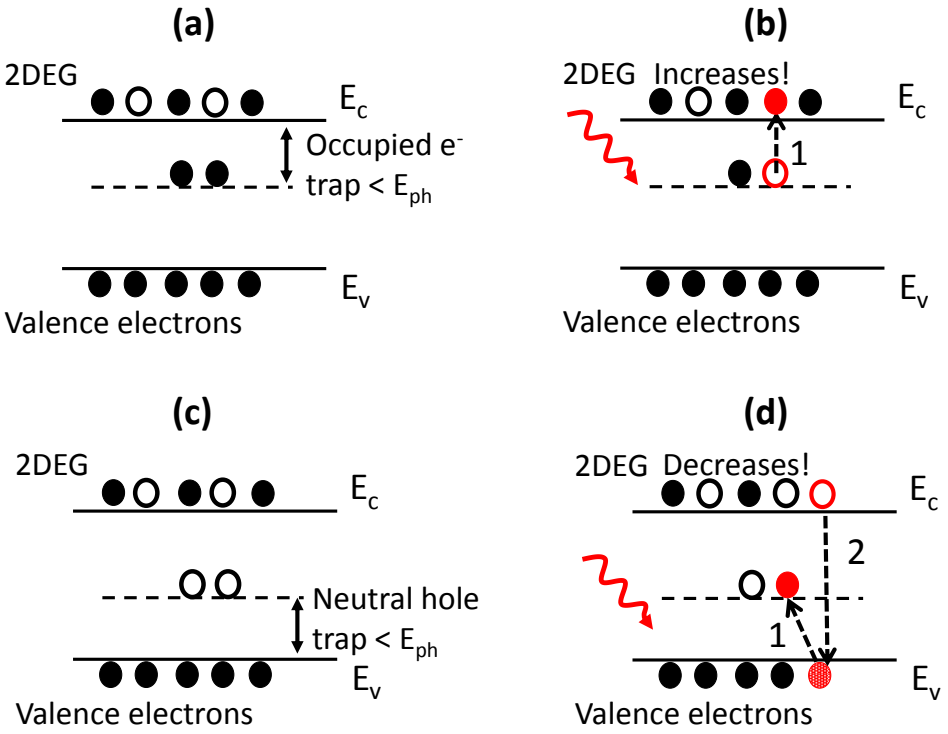


Figure 8-3. The carrier dynamics of the photoionization spectroscopy technique are illustrated. (a) Device is under dark conditions with trapped electrons. (b) Device in "(a)" is illuminated with photons of energy greater than the electron trap. (c) Device is under dark conditions with neutral hole traps. (d) Device in "(c)" is illuminated with photons of energy greater than the neutral hole trap.

the reversal of drain current collapse in GaN devices [82–86]. The carrier dynamics of photoionization spectroscopy are shown in Figure 8-3. For a device under dark conditions with occupied electron traps of activation energy E_A (Figure 8-3-a), optical illumination from an incident photon with an energy equal to or greater than E_A results in the emission of the trapped electron to the GaN conduction band. The photoionization of the trapped electron is shown as "1" in Figure 8-3-b. With the assumption that the time required to traverse the HEMT channel is shorter than the trap capture time, the trap state at E_A below the conduction band is continuously emptied to a non-equilibrium steady-state, proportional to the density of incident photon flux. Hence, the electrons residing at multiple traps with activation energy lower than the incident photon energy contribute to the 2DEG and increase the drain current during the optical illumination.

Assuming that the carrier mobility remains unchanged and light intensity is high enough to achieve complete photoionization of all trap locations for the given photon energy, the ratio of the total trap density to the 2DEG is obtained by comparing the increased drain current under illumination to the dark current. For a small incremental change in the incoming photon wavelength, the areal trap density for the given energy is approximated by,

$$\frac{D_t(h\nu_f)}{N_{2DEG}} = \frac{I(h\nu_f) - I(h\nu_i)}{I_{dark}} \frac{1}{h\nu_f - h\nu_i}, \tag{8–1}$$

where ν_f and ν_i are the final and initial photon frequencies, I_{dark} is the dark current, and $I(h\nu_f)$ and $I(h\nu_i)$ are the stabilized drain current under illumination with final and initial photon frequencies.

While the increase in the drain current is observed due to the photoionized trapped electrons, neutral hole traps with a smaller activation energy than the incoming photon energy decreases the drain current. This phenomenon is called “optical quenching effect” and is observed in GaN based devices, especially under red and yellow band illumination [241, 242]. The fundamentals of optical quenching is shown in Figs. 8-3c and 8-3d. When a device is exposed to optical illumination with an incoming photon energy greater than the activation energy of a neutral hole trap, an electron in the valence band is excited to the neutral hole trap location, shown as “1” in Figure 8-3-d. Then, the resulting hole in the valence band recombines with an electron from the 2DEG and decreases the current (“2” in Figure 8-3-d). Therefore, not only electron traps but also hole traps can be characterized using the photoionization spectroscopy technique. Consequently by using a set of different subbandgap light wavelengths, the energy distribution of the areal electron and hole trap densities may be characterized.

8.4 Experimental Setup

To investigate the trap generation behavior in AlGaN/GaN heterostructures, the experimental setup shown in Figure 8-4 is designed to conduct trap generation and characterization measurement, via electrical stressing and photoionization spectroscopy

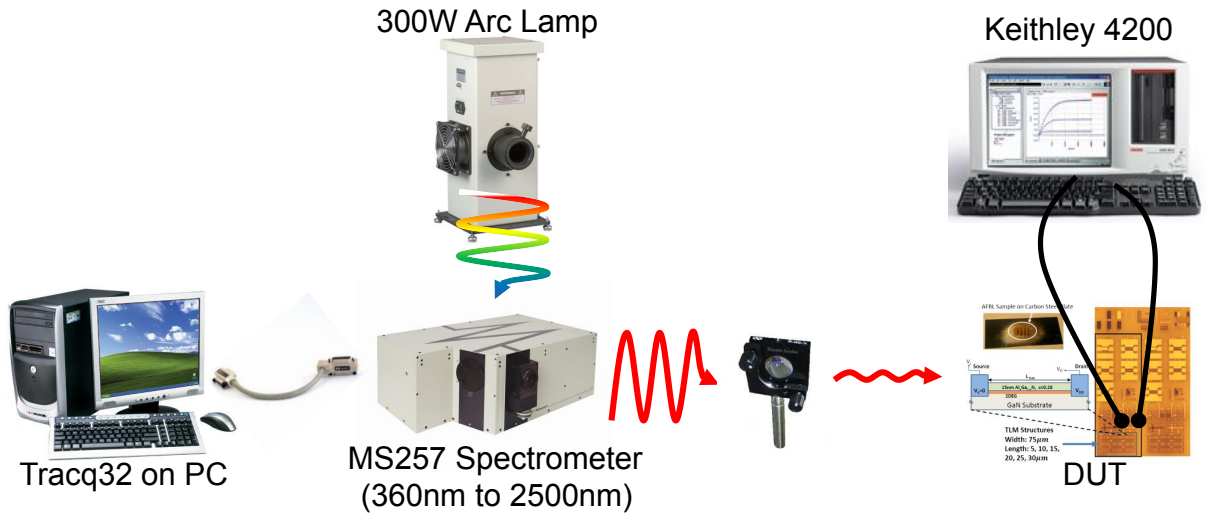


Figure 8-4. The experimental setup used in this study is shown. The broad spectrum radiation from the arc lamp is monochromated by using a computer controlled Oriel MS257 spectrometer. Then, the monochromated light is routed and focused on the device-under-test (DUT). The DUT is electrically probed by the Keithley 4200-SCS semiconductor characterization system for both trap characterization and trap generation experiments.

(PS) techniques respectively. For the characterization of the traps, a continuous wavelength sweep monochromator is used. This ensures that a sufficient resolution in trap energy distribution can be achieved. The continuous wavelength sweep capability is realized by using a 300W arc lamp coupled with a computer controlled MS257 Oriel spectrometer located at the University of Florida Microfabritech facility. The beam at the monochromator aperture is collimated and focused using two optical neutral density lenses, while it is shone directly on the GaN sample in the 4-point bending apparatus using a metal mirror. The size of the beam spot on the sample is as small as approximately 3mm by 1mm, which in turn increases the photon flux density incident on the DUT. The effects of the increased incident photon flux density is shown in Figure 8-5. With a focused beam, the increased flux of incident photons reduces the transient time to reach the non-equilibrium steady-state, while increasing the level of the steady-state current due to a larger percentage of the photoionized electrons at this trap energy.

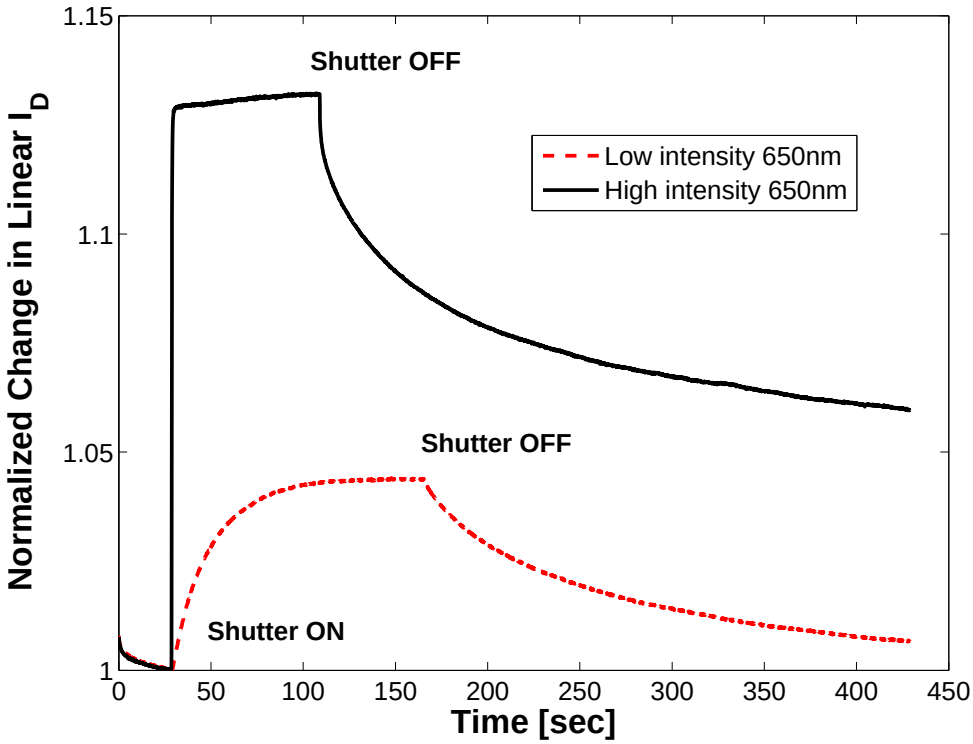


Figure 8-5. The linear drain current response from a 30μm long TLM device is shown for both with and without focusing lenses (i.e. high and low intensity) 650nm illumination.

By avoiding an optical fiber light guide, optical intensity loss is minimized. During PS experiments, a Keithley 4200 system is used in the sampling mode to measure the linear drain current as a function of time. The measurement speed, filtering, and delay settings are optimized to achieve a desirable tradeoff between the signal to noise ratio and the sampling rate to capture fast trapping-detrapping behaviors. The drain voltage is set to a value within the linear operation regime of the device during PS measurements to ensure that the current forced through the device does not cause further degradation or additional trap generation.

During the wavelength sweep, a UV-VIS band grating with 1200 lines/mm and a blazing wavelength of 350nm is used to decompose the wavelengths of the radiation from the arc lamp. In order to avoid additional electron-hole pair photogeneration, subbandgap illumination is performed, setting the shortest photon wavelength to the

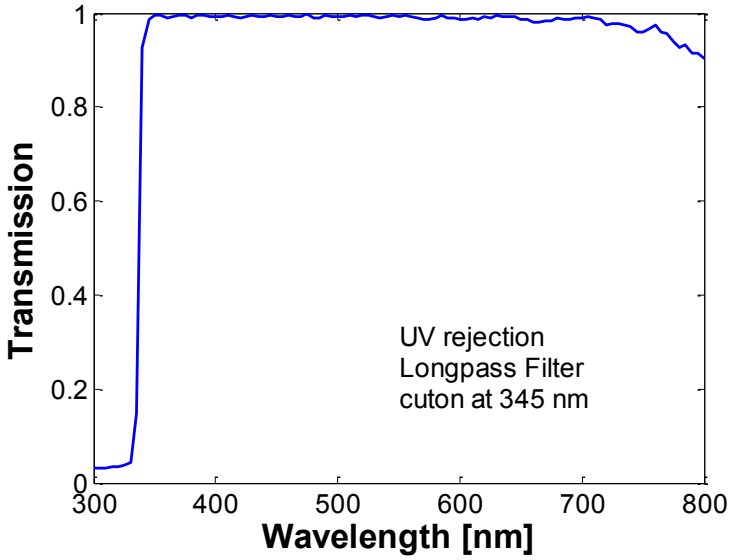


Figure 8-6. The transmission characteristics of the UV-rejection higher order filter is shown.

365-360nm range. To determine the value of the longest incident photon wavelength, the effects of higher-order stray light is taken into consideration. For a given incident angle of the arc lamp radiation onto a grating preset to a wavelength λ, the higher-order photons with wavelengths of λ/n leak from the grating at a reduced intensity. For example, a grating set for 680nm red illumination would also pass UV components at 340nm, 170nm, etc. with ever decreasing intensity. If the higher-order wavelengths are not blocked, this would lead to multiwavelength illumination instead of the desired set wavelength and result in an erroneous trap characterization effort, especially in the case of higher-order above GaN bandgap photon energies. To minimize the experimental errors due to higher-order wavelengths and to increase the measurement accuracy, we have placed a UV-rejection longpass filter with a cut-on wavelength at 345nm. The measured transmission characteristics of the UV-rejection filter is shown in Figure 8-6. By noting the longest filtered wavelength (340nm), we set the maximum of the PS wavelength sweep to 680nm. This enabled us to perform a higher-order free illumination from 680nm to 360nm without any attenuation caused by the UV-rejection filter.

In Section 8.3, we have discussed the physics of the photoionization spectroscopy technique. It was emphasized that Eq. 8–1 is valid once the non-equilibrium steady-state photoexcited drain current is reached for a given wavelength. The carrier dynamics shown in Figure 8-3 are modified with the incident photon count per unit area at a given time. For example at a constant temperature, a device in the dark with trapped electrons is at thermal equilibrium with equilibrium carrier concentration determined by trapping and emission of electrons between the conduction band and the electron trap level. Once the device is illuminated, the incoming photons change the equilibrium condition to a non-equilibrium state with a more favorable emission rate than trapping. Next, the emitted electrons increase the 2DEG density and modify the equilibrium to a more favorable trapping rate. With the interplay between the increasing detrapping rate by the incident photons and the instantaneous increase in the trapping rate due to higher 2DEG density, the device reaches to a steady-state condition under the non-equilibrium optical illumination condition. For an incomplete photoionization case, the percentage of detrapped electrons at the non-equilibrium steady-state is incident photon flux dependent. Therefore to interpret the measured PS data correctly, the relative incident photon flux for each wavelength is necessary.

The optical intensity spectrum is measured using a combination of Oriel 70124 pyroelectric detector, Spectra-650 colorimeter, and an Oriel 77346 photomultiplier tube. The reason to use multiple detectors is that none of the three covers the wavelength range of interest completely. The pyroelectric detector has flat responsivity for wavelengths longer than the cut-off at 580nm. Hence, the pyroelectric detector is used to verify the reading from the Spectra-650 colorimeter for wavelengths longer than 580nm (Figure 8-7). While the Spectra-650 is verified above 580nm, it has a cut-off lower than 380nm. Finally, by using the Oriel 77346 photomultiplier tube and the colorimeter for the 400nm to 360nm range , we were able to obtain the optical intensity for the entire range of interest. The relative optical intensity is obtained through

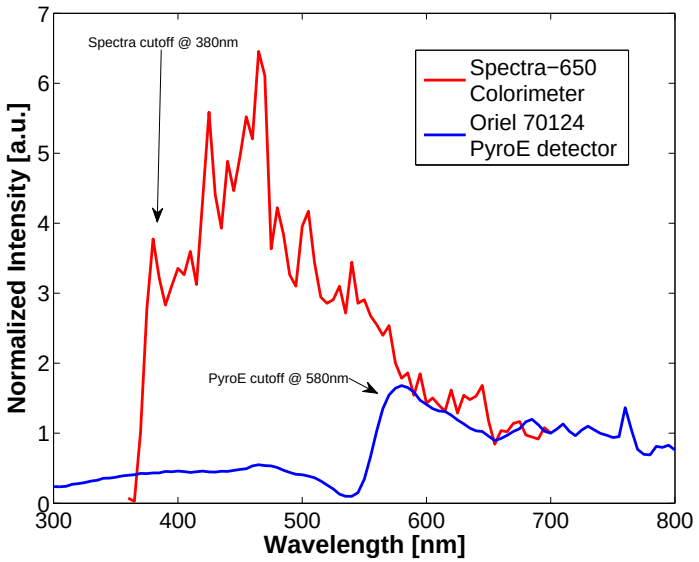


Figure 8-7. Cross calibration curves between Spectra-650 colorimeter and Oriel 70124 pyroelectric detector.

the ratio of the entire intensity spectrum to the near bandgap illumination intensity at 360nm, since at this wavelength all the subbandgap traps energy levels are excited. Then, by normalizing the relative optical intensity with the incident photon energy, we have obtained the relative photon flux for each wavelength as shown in Figure 8-8.

Since the incident photon flux is different for different wavelengths, the measured percent change in linear drain current would be intensity dependent for an incomplete photoionization assumption. The non-zero slope of the photoexcited current versus time at the non-equilibrium quasi-steady-state of the high intensity illumination shown in Figure 8-5 suggests that our setup partially photoionizes the trapped electrons. To account for the wavelength dependent photon flux, Eq. (8–1) was modified to obtain a photon flux-normalized relative areal trap density distribution given by,

$$\frac{RD_t(h\nu_f)}{N_{2DEG}} = \frac{I(h\nu_f) - I(h\nu_i)}{I_{dark}} \frac{1}{(h\nu_f - h\nu_i)C_f}, \tag{8–2}$$

where C_f is the relative photon flux computed from the ratio of the entire intensity spectrum to the near bandgap illumination intensity at 360nm per photon energy. Since

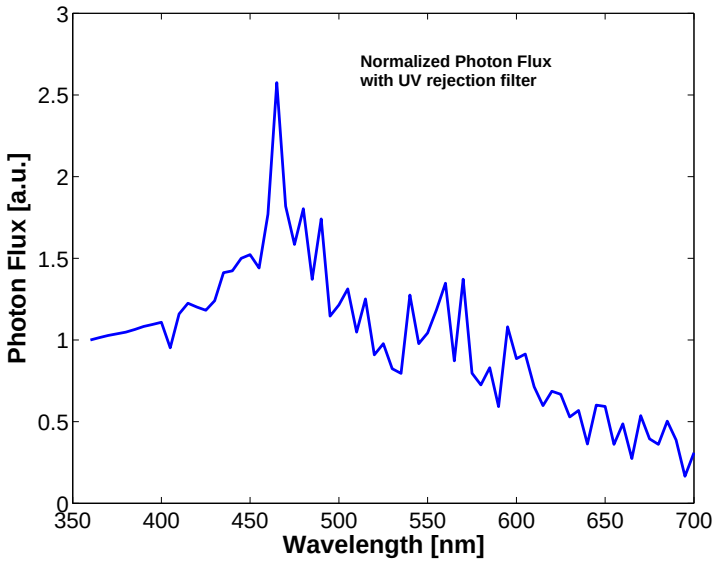


Figure 8-8. The normalized photon flux with respect to 360nm illumination vs. incident photon wavelength is shown.

the exposure time is the same for each wavelength, normalization by photon flux is the same as normalization by photon count. The term $1/[(h\nu_f - h\nu_i)C_f]$, a density normalization coefficient, represents the normalized effective trap density per energy at energy $h\nu_f$ corresponding to the drain current change per increment in wavelength. For a wavelength sweep from 650nm to 360nm with decrements of 5nm, the density normalization coefficient is shown with respect to both photon wavelength and energy in Figure 8-9. Hence, the percent change in the linear drain current obtained from the PS measurements are calibrated with the density normalization coefficient to normalize the measured areal trap density at each energy. Although incomplete photoionization of traps hinders the capability of this method to measure the energy distribution of the absolute areal trap density ($D_t(h\nu)$), under the same illumination conditions, the relative changes of the measured normalized trap density ($RD_t(h\nu)$) before and after device degradation experiments can be monitored reliably, providing key insights into the type of trap generation responsible for device degradation.

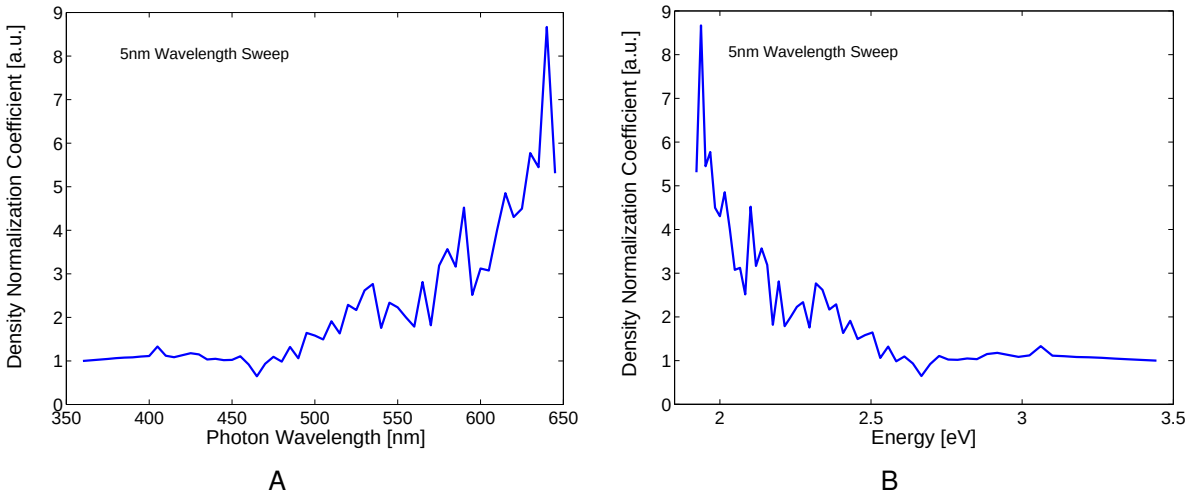


Figure 8-9. The density normalization coefficient is shown for different photon (a) wavelengths and (b) energies for a wavelength sweep of 5nm steps.

While the Keithley 4200-SCS system is used for characterization, the high power state electrical stressing of the devices were performed with the custom built electrical stress and characterization system located at the Nanoscale Research Facility at the University of Florida [86]. This high power system is capable of applying $\pm$60V to all three terminals of the device while passing several hundred milliamperes for a desired period of time and stress condition, such as step stress, constant stress, step-recovery stress, or pulse-recovery stress.

8.5 Results and Discussion

In order to avoid self-heating effects, the photoinduced drain current transients are measured at the low V_{DS} region, when the device is in the linear operation regime. Another motivation for conducting the photoionization spectroscopy experiments at low drain biases is to avoid further trap generation during characterization of the existing traps. Figure 8-10 shows the current-voltage data measured from 30μm TLM structures from Si and SiC substrate samples. The close agreement between the pre (black) and post (red) photoionization spectroscopy I-V curves under dark condition indicates that the trap distribution of the measured devices are not altered during device

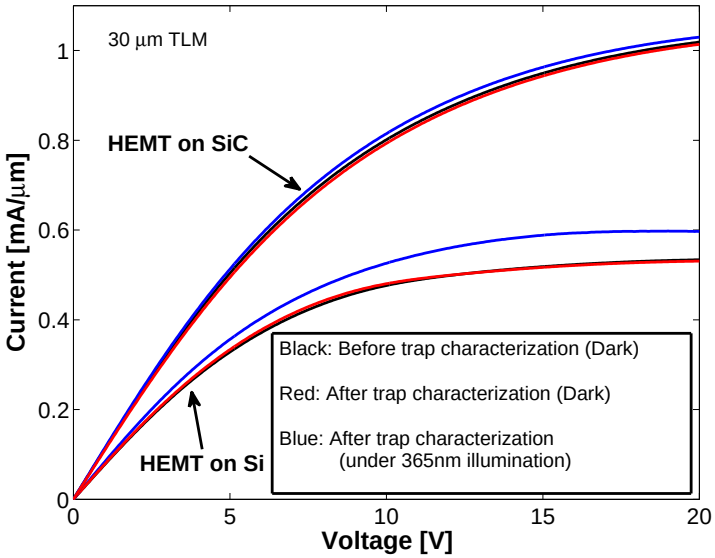


Figure 8-10. The current-voltage curves are shown for 30 μm TLM structures on Si and SiC substrates. Measured dark current before (black) and after (red) trap characterization experiments are shown. The blue curves denote I-V obtained under 365nm illumination after PS measurements.

characterization. It is immediately seen that the AlGaN/GaN HEMT on SiC devices have larger drive current which is attributed to the threshold voltage based difference in the 2DEG density as shown in Figure 8-2-b. A much lower sensitivity to near bandgap illumination is observed in SiC based device compared to Si devices, which is expected due to the larger trap density of the Si devices.

A typical photoionization spectroscopy data measured from a 5μm TLM site on Si substrate is shown in Figure 8-11. The wavelength is swept from 650nm to 360nm with decrements of 5nm and an exposure time of 60 seconds at each wavelength. The shutter is turned off after 60 seconds of band gap illumination at 360nm. After the light is turned off, the linear drain current decreases immediately due to the recombination of the small number of photoexcited excess electron-hole pairs with 360nm illumination [66]. The initial sharp drop due to recombination is a very fast transient due to the very low intrinsic carrier concentration of the GaN ($\sim 10^{-10} cm^{-3}$) [32]. After the initial sharp drop, the linear drain current decays exponentially.

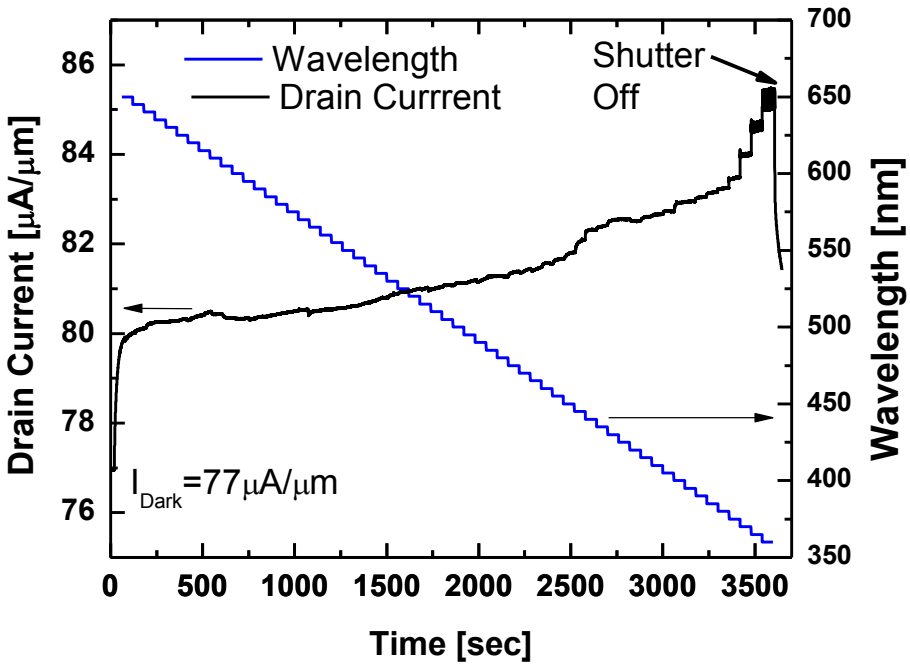


Figure 8-11. A typical photoionization spectroscopy data from a 5μm TLM site on Si substrate is shown. For a wavelength sweep of 650nm to 360nm with decrements of 5nm and exposure time of 60 seconds, the photoexcited drain current evolution with changing wavelengths over time is presented.

The evolution of the near steady-state non-equilibrium photoexcited drain current with time is used to obtain the relative changes between increments in the photon energy. Figure 8-12 shows the extracted $RD_t(h\nu)$ for TLM structures on Si and SiC substrates, respectively. The total number of traps in SiC substrate devices is found to be smaller compared to the Si substrate devices. The main difference between SiC and Si substrate devices is observed at the near bandgap trap densities. The near band gap trap density in SiC based devices is less the half of that in Si substrate. It has been shown that the near bandgap traps are attributed to structural defects in GaN [59, 92]. Since SiC has a closer lattice constant match with GaN compared to Si (111) substrate, a lower structural defect density is expected, reducing the photosensitivity of the drain current to near band gap illumination.

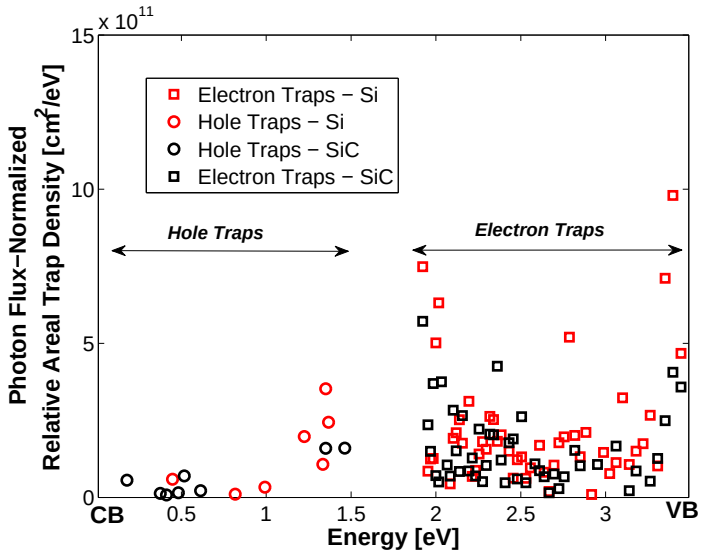


Figure 8-12. Extracted photon flux-normalized relative areal trap density is shown for a 30μm long TLM on SiC and a 5μm long TLM on Si substrate. The dark condition 2DEG densities are approximated using the CV data from the CV-dot devices on each substrate types.

After characterization of the photon flux-normalized relative areal trap density distribution on virgin TLM structures on Si and SiC substrates, we have demonstrated the repeatability of the photoionization spectroscopy method using a virgin 100nm off-center gate HEMT on SiC substrate. To achieve this, we have conducted two consecutive PS experiments on the same virgin device. Between each PS run, the device was kept in dark and the terminal currents were monitored to ensure that the device returns to pre-PS condition. During the PS measurements, the drain voltage was kept within the linear regime at 0.2V. In order to ensure that the PS signal was obtained from the gated portion of the channel, the gate bias was adjusted such that the gated portion of the channel dominates the total source-to-drain resistance. Using the CV data, the gate voltage for the PS experiments was chosen to be -3.3V, for which the 2DEG density under the gate is $1 \cdot 10^{10} cm^{-2}$. Hence, the 0.1μm-long gated portion of the channel has 3 orders of magnitude lower electron density compared to the 3.9μm-long ungated part of the channel. Therefore, the changes in the total channel resistance are solely due to the photoionization of the trapped electrons under the gate. The data

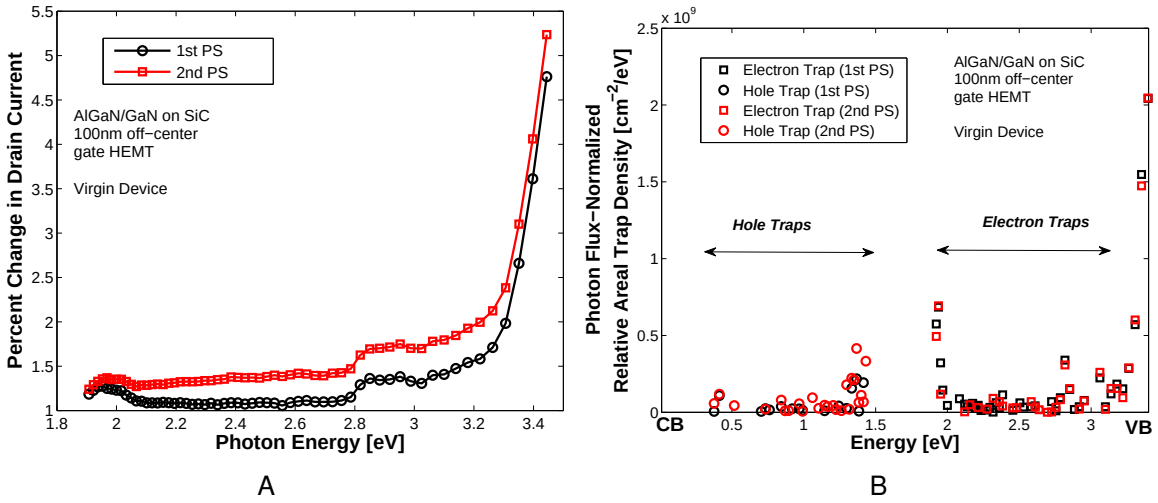


Figure 8-13. The repeatability of the PS experiment is demonstrated on a virgin 100nm off-center gate HEMT on SiC substrate. (a) The percent change in drain current is shown for two consecutive PS runs. (b) The extracted photon flux-normalized relative areal trap density is shown using the data shown in left.

obtained from repeated PS measurements is shown in Figure 8-13. The error in the percent change in photoionized drain current for repeated measurements is found to be on the order of 0.5%. Hence, a good agreement is observed in the energy relative distribution of extracted areal trap density for consecutive measurements.

Next, the effects of OFF-state electrical gate stressing on the HEMT trap characteristics are investigated using photoionization spectroscopy. A 100nm off-center HEMT on SiC device was controllably degraded by electrical step stressing of the gate from V_G=-5V with decrements of 1V for a period of 70 seconds each. During the OFF-state gate stress, both drain and source terminals were grounded. The device is subjected to OFF-state stress twice to investigate both initial and further degradation effects. In the first run, similar to the method in [20], the gate current density is monitored while stepping the reverse gate bias until a sudden jump is observed. Once the drastic sudden increase in gate current is detected, the experiment is automatically stopped before the device is further degraded. For the second run, the device is further stressed until the

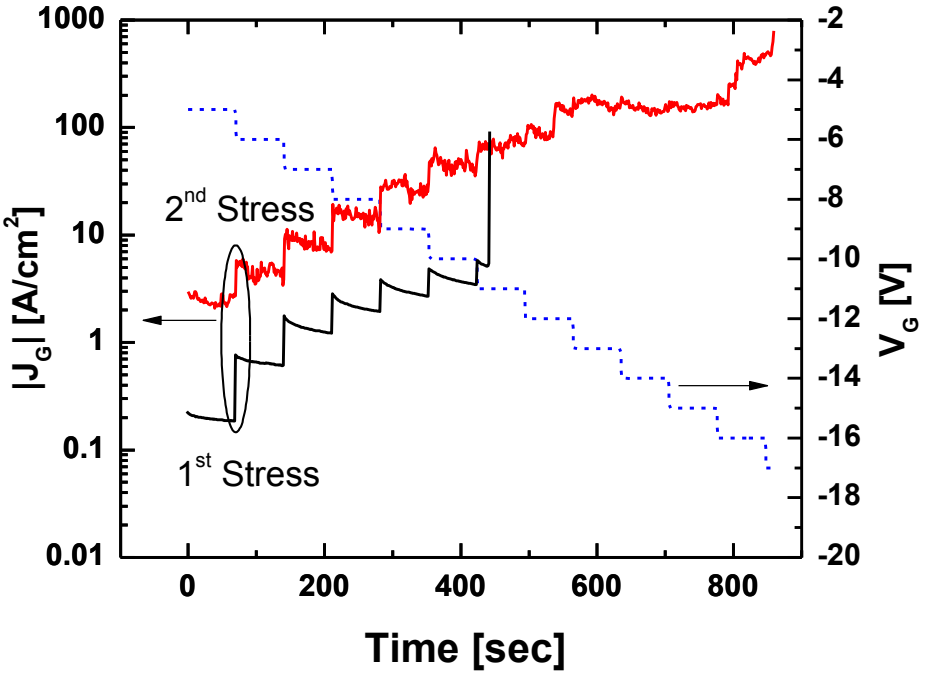


Figure 8-14. Data collected from two consecutive runs of OFF-state electrical gate stress experiments is shown for a 100nm off-center gate HEMT on SiC substrate. During the 1st stress experiment, the gate current density jump is observed at V_{crit}=-11V. For the 2nd run of the OFF-state stressing, the gate leakage is found to be much noisier and one order of magnitude larger.

gate current density hits the limit of 800A/cm^2. The electrical OFF-state stressing data is shown in Figure 8-14. The critical voltage for which the sudden jump in the gate current was observed is found to be V_{crit}=-11V, comparable to the results shown in [243]. For the second stress experiment, the gate current is found to be an order of magnitude larger with a significant noise component.

The pre-stress and post-stress gate current density and width normalized drain current of the device are shown in Figure 8-15 and Figure 8-16. Compared to the virgin device, the gate leakage has increased by an order of magnitude after the first stress run. While the changes in the gate current are significant, the difference between the pre-stress and post-stress linear drain currents is negligible. On the other

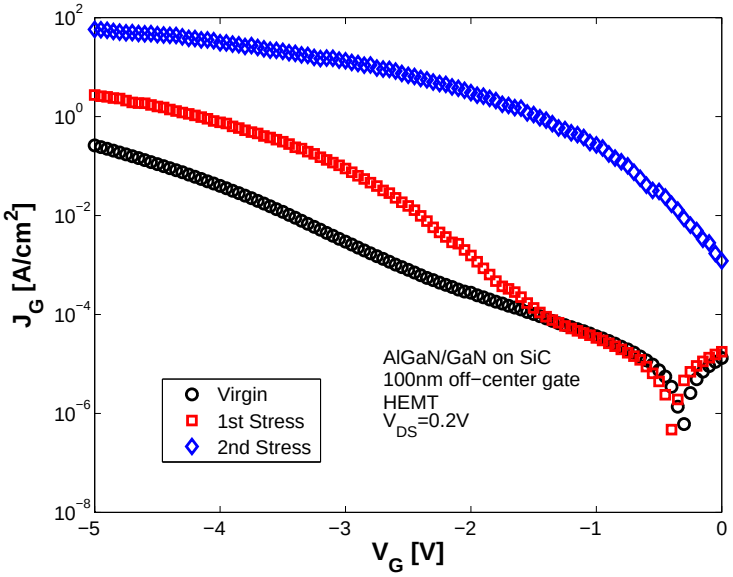


Figure 8-15. Gate current density is shown for pre-stress and post-OFF-state electrical stress conditions. With second stress run, the gate leakage has increased more than two orders of magnitude.

hand the second gate stressing resulted in a significant degradation of the device characteristics. The gate leakage has increased by more than 2 orders of magnitude, becoming comparable to the drain current. While the 2DEG density was found to be unchanged for the given degradation condition, a decrease in the drain current is observed for a given gate bias, indicating a mobility degradation. Yet, at large gate voltage range (above -2.5V), the parasitic R_{SD} dominates the channel resistance and the effect of reduced mobility becomes negligible.

The trap characteristics of the HEMT device are experimentally extracted both before and after OFF-state stress, using the photoionization spectroscopy method outlined previously. The pre-stress and post-stress photon flux-normalized relative areal trap densities are shown in Figure 8-17. Compared to the virgin device, near bandgap electron trap densities have increased during each run of electrical OFF-state stressing. The increase in the near bandgap traps indicates an increased structural defect density

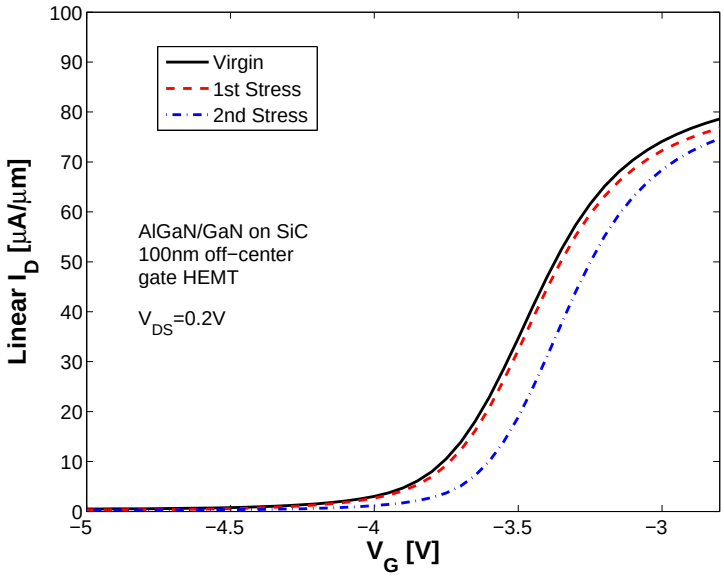


Figure 8-16. Width normalized drain current is shown for pre-stress and post-OFF-state electrical stress conditions. While the first stressing does not change the drain current much, further stressed device experiences a reduction in 2DEG mobility.

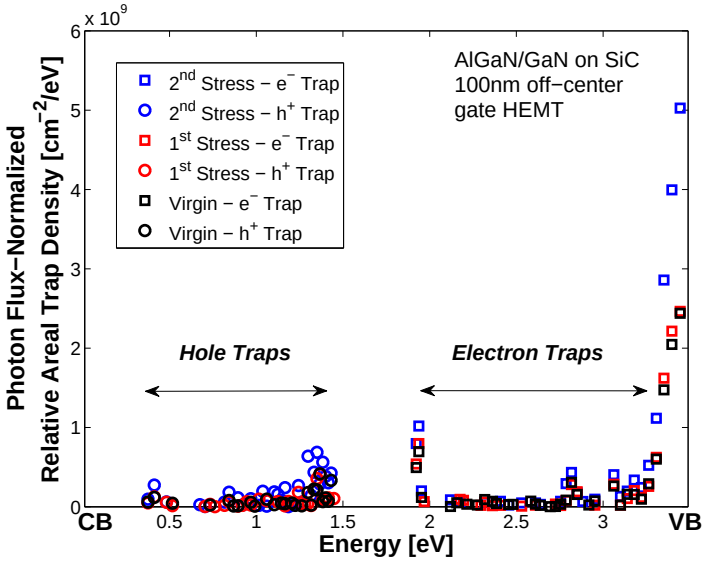


Figure 8-17. Extracted photon flux-normalized relative electron and hole areal trap densities are shown for before and after the OFF-state stress. A significant increase in the structural defect density (near bandgap traps) is observed.

[59, 92]. For a gate voltage stress at and beyond the V_{crit}, this result is in accordance with the “crack” formation hypothesis due to the inverse piezoelectric effect [20].

8.6 Summary

Using the photoionization spectroscopy technique, we have devised an experimental baseline to measure and characterize traps in AlGaN/GaN based devices. The details of the method and the experimental setup are provided. The energy relative distribution of the areal electron and hole trap densities from TLM and HEMT devices with both Si and SiC substrates are extracted. It is found that the SiC based devices have much lower trap densities around the GaN bandgap, which is attributed to the structural defects in the GaN. For HEMT devices subjected to electrical OFF-state gate stressing, it is found that the electron traps related to the structural GaN defects has increased. This supports the previous findings in the literature on the hypothesis of crack formation at the drain edge of the gate due to inverse piezoelectric effect.

CHAPTER 9
SUMMARY AND RECOMMENDATION FOR FUTURE WORK

9.1 Research Summary

The main focus of this research is to understand the effects of mechanical strain on silicon MOS device performance enhancement as the transverse and longitudinal device dimensions scale down and on the reliability of AlGaN/GaN HEMT devices.

Physical insights on the experimentally observed comparable electron mobility in (100) and (110) sidewall undoped 20nm thick silicon FinFETs with high-κ/TiN gate are provided. (100) and (110) sidewall silicon FinFETs with boron and phosphorus doped channels, SiO_2/poly-Si gate stacks, and thicker fins were fabricated to explore the potential underlying reasons such as undoped channel, metal-gate induced stress, and volume-inversion. Comparable electron mobility on (100) and (110) FinFETs was observed from all of these test structures. Using a modified self-consistent double-gate FinFET simulator, it was shown for (110) FinFET that the effect of the heavier nonparabolic confinement mass of the Δ_2 valley leads to a much reduced valley energy, making Δ_2 the ground valley. Hence, inversion electrons mostly occupy the Δ_2 valley for both surface orientations and thus, a comparable average electron transport effective mass is obtained.

Following the discussion on the electron transport in double-gate FinFETs, the width dependence of the strain response of p-type tri-gate FinFETs was investigated. Using a 4-point wafer bending jig, low-field piezoresistance coefficients of tri-gate FinFETs with varying widths from 500nm to 20nm were extracted at high overdrive conditions. It was found that for the fin width range investigated, volume-inversion effects were negligible. The extracted π-coefficients of tri-gate FinFETs were accurately modeled using a weighted conductance average of the top and sidewall surface bulk piezoresistive coefficients .

With further scaling of the silicon FinFETs both in transverse and longitudinal directions, gate-all-around nanowire transistors operating at near ballistic limit were considered. The strain enhancement of p-type ballistic silicon nanowire MOSFETs was studied using $sp^3d^5s^*$ basis nearest-neighbor tight-binding simulations. By coupling the tight-binding bandstructure to a compact electrostatic model, high computational efficiency was achieved while providing similar trends to self-consistent tight-binding simulations. The impact of nanowire channel direction and cross section size on the strain enhancement of ballistic hole transport was explained by the strain-induced modification of the 1D nanowire valence band density-of-states. Further insights were provided for future p-type high-performance silicon nanowire logic devices.

In order to understand the complete scheme of strain effects in n-type silicon MOS devices, a physics based investigation of the strain enhancement of surface roughness limited mobility was conducted. Experimental evidence of a possible strain-induced morphology change was obtained through electrical characterization of strained nMOS devices at low temperatures and AFM measurements of silicon samples under external mechanical bending. The *ad hoc* modification of surface morphology was validated by the agreement between electrical and materials characterization and the transport simulations.

To link the strain effects on low-field linear current to the quasi-ballistic saturation current, a unified carrier transport model applicable to all range of MOSFET operation conditions was developed using an updated one-flux theory. A new surface-potential based MOSFET channel transmission expression was developed with the inclusion of high-field and carrier confinement effects. The ballisticity of the transistor was redefined using the transmission expression and the average carrier drift velocity. The transmission at high-field and strong inversion conditions was shown to be dependent on the “diffusibility” of the channel, which links the drift-diffusion dominant low-field transport to high-field quasi-ballistic saturation current in one unified model. By using

our unified model, MOSFET channel transmission characteristics were determined under given gate and drain biases. This in turn was used to understand the strain modification of the average carrier drift velocity and channel transmission, through an interplay of strain-induced modification of carrier effective mass and scattering rates.

To understand the impact of mechanical strain on AlGaN/GaN HEMT trap generation, an experimental method to obtain the photon flux-normalized relative areal trap density using photoionization spectroscopy technique was developed. Using this method, the trap characteristics were extracted for both ungated transmission line module (TLM) and gated HEMT devices from both Si and SiC substrates. It was found for SiC substrate devices that the built-in structural defect density was much lower compared to AlGaN/GaN HEMT devices on Si substrate due to a closer lattice constant match between SiC and GaN. The effect of step electrical OFF-state gate stressing was investigated for HEMT on SiC devices. For repeated gate stressing experiments on the same HEMT device, it was shown that the structural defect density increases at the end of each successive degradation experiment. The structural defect formation was linked to an increase in the biaxial tension in the AlGaN barrier through the inverse piezoelectric effect at high reverse biases, similar to observations from other research groups.

9.2 Recommendation for Future Work

As we have discussed, strained silicon nanowire field-effect-transistors are expected to be introduced into high-volume manufacturing at sub-10nm CMOS technology nodes. In order to further understand the advantages and issues related to this technology, a size and orientation dependent experimental study of strain effects on electron and hole transport in silicon nanowires is strongly recommended. In addition to these, a finite element method (FEM) based investigation of an ideal process-induced stressor scheme for NWFETs is required to exploit the benefits of mechanical strain at device integration level. The cumulative findings of these efforts may translate into a more successful

and cost-effective process integration technology development of Si NWFET based high-volume manufacturing IC technology.

In order to improve the understanding of the root causes of AlGaN/GaN HEMT reliability issues, a simulation based modeling of the OFF-state stress gate leakage and device degradation is recommended. To achieve this, measured changes in the energy distribution of the areal trap densities due to the device degradation experiments can be incorporated to full-scale device simulators such as Florida Object Oriented Reliability Simulator (FLOORS). As a result, further insights on additional degradation physics may be obtained by simulating different device structure and material choices for AlGaN/GaN HEMT devices. Based on these simulations, devices showing improved degradation characteristics may be fabricated for experimental verification. Consequently, these verified device structures may be used to improve the reliability of AlGaN/GaN HEMT technology.

REFERENCES

[1] S. Thompson, N. Anand, M. Armstrong, C. Auth, B. Arcot, M. Alavi, P. Bai, J. Bielefeld, R. Bigwood, J. Brandenburg, M. Buehler, S. Cea, V. Chikarmane, C. Choi, R. Frankovic, T. Ghani, G. Glass, W. Han, T. Hoffmann, M. Hussein, P. Jacob, A. Jain, C. Jan, S. Joshi, C. Kenyon, J. Klaus, S. Klopcic, J. Luce, Z. Ma, B. Mcintyre, K. Mistry, A. Murthy, P. Nguyen, H. Pearson, T. Sandford, R. Schweinfurth, R. Shaheed, S. Sivakumar, M. Taylor, B. Tufts, C. Wallace, P. Wang, C. Weber, and M. Bohr, "A 90 nm logic technology featuring 50 nm strained silicon channel transistors, 7 layers of cu interconnects, low k ild, and 1 m2 sram cell," in *Electron Devices Meeting, 2002. IEDM '02. Digest. International*, 2002, pp. 61–64.

[2] Y. Sun, S. E. Thompson, and T. Nishida, "Physics of strain effects in semiconductors and metal-oxide-semiconductor field-effect transistors," *Journal of Applied Physics*, vol. 101, no. 10, pp. 104503, 2007.

[3] M. Chowdhury, V. Trivedi, J. Fossum, and L. Mathew, "Carrier mobility/transport in undoped-utb dg finfets," *Electron Devices, IEEE Transactions on*, vol. 54, no. 5, pp. 1125–1131, May 2007.

[4] B. Yu, L. Chang, S. Ahmed, H. Wang, S. Bell, C.-Y. Yang, C. Tabery, C. Ho, Q. Xiang, T.-J. King, J. Bokor, C. Hu, M.-R. Lin, and D. Kyser, "Finfet scaling to 10 nm gate length," in *Electron Devices Meeting, 2002. IEDM '02. Digest. International*, 2002, pp. 251–254.

[5] S. Suthram, M. Hussain, H. Harris, C. Smith, H.-H. Cheng, R. Jammy, and S. Thompson, "Comparison of uniaxial wafer bending and contact-etch-stop-liner stress induced performance enhancement on double-gate finfets," *Electron Device Letters, IEEE*, vol. 29, no. 5, pp. 480–482, May 2008.

[6] H. Lee, L.-E. Yu, S.-W. Ryu, J.-W. Han, K. Jeon, D.-Y. Jang, K.-H. Kim, J. Lee, J.-H. Kim, S. C. Jeon, G. S. Lee, J. S. Oh, Y. C. Park, W. H. Bae, H. M. Lee, J. M. Yang, J. J. Yoo, S. I. Kim, and Y.-K. Choi, "Sub-5nm all-around gate finfet for ultimate scaling," in *VLSI Technology, 2006. Digest of Technical Papers. 2006 Symposium on*, 0-0 2006, pp. 58–59.

[7] M. Bohr, "The evolution of scaling from the homogeneous era to the heterogeneous era," in *Electron Devices Meeting (IEDM), 2011 IEEE International*, dec. 2011, pp. 1.1.1 –1.1.6.

[8] G. Moore, "No exponential is forever: but "forever" can be delayed! [semiconductor industry]," in *Solid-State Circuits Conference, 2003. Digest of Technical Papers. ISSCC. 2003 IEEE International*, 2003, pp. 20 – 23 vol.1.

[9] K. Natori, "New solution to high-field transport in semiconductors: I. elastic scattering without energy relaxation," *Japanese Journal of Applied Physics*, vol. 48, no. 3, pp. 034503, 2009.

[10] K. Natori, "New solution to high-field transport in semiconductors: Ii. velocity saturation and ballistic transmission," *Japanese Journal of Applied Physics*, vol. 48, no. 3, pp. 034504, 2009.

[11] V. Barral, T. Poiroux, F. Rochette, M. Vinet, S. Barraud, O. Faynot, L. Tosti, F. Andrieu, M. Casse, B. Previtali, R. Ritzenthaler, P. Grosgeorges, E. Bernard, G. LeCarval, D. Munteanu, J. Autran, and S. Deleonibus, "Will strain be useful for 10nm quasi-ballistic fdsoi devices? an experimental study," *VLSI Technology, 2007 IEEE Symposium on*, pp. 128–129, June 2007.

[12] K. Natori, "Ballistic metal oxide semiconductor field effect transistor," *Journal of Applied Physics*, vol. 76, no. 8, pp. 4879 –4890, oct 1994.

[13] U. Mishra, P. Parikh, and Y.-F. Wu, "Algan/gan hemts-an overview of device operation and applications," *Proceedings of the IEEE*, vol. 90, no. 6, pp. 1022 – 1031, jun 2002.

[14] M. A. Khan, A. Bhattarai, J. N. Kuznia, and D. T. Olson, "High electron mobility transistor based on a gan-al[sub x]ga[sub 1 - x]n heterojunction," *Applied Physics Letters*, vol. 63, no. 9, pp. 1214–1215, 1993.

[15] R. Trew, "High-frequency solid-state electronic devices," *Electron Devices, IEEE Transactions on*, vol. 52, no. 5, pp. 638 – 649, may 2005.

[16] S. Adachi, *Handbook of Physical Properties of Semiconductors*, Kluwer Acadmic Publishers, Norwell, MA., USA, 2004.

[17] Y.-F. Wu, A. Saxler, M. Moore, R. Smith, S. Sheppard, P. Chavarkar, T. Wisleder, U. Mishra, and P. Parikh, "30-w/mm gan hemts by field plate optimization," *Electron Device Letters, IEEE*, vol. 25, no. 3, pp. 117 – 119, march 2004.

[18] O. Ambacher, J. Smart, J. Shealy, N. Weimann, K. Chu, M. Murphy, W. Schaff, L. Eastman, R. Dimitrov, L. Wittmer, and Others, "Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N-and Ga-face AlGaN/GaN heterostructures," *Journal of Applied Physics*, vol. 85, pp. 3222, 1999.

[19] J. P. Ibbetson, P. T. Fini, K. D. Ness, S. P. DenBaars, J. S. Speck, and U. K. Mishra, "Polarization effects, surface states, and the source of electrons in algan/gan heterostructure field effect transistors," *Applied Physics Letters*, vol. 77, no. 2, pp. 250–252, 2000.

[20] J. Joh and J. del Alamo, "Critical voltage for electrical degradation of gan high-electron mobility transistors," *Electron Device Letters, IEEE*, vol. 29, no. 4, pp. 287 –289, april 2008.

[21] G. Meneghesso, G. Verzellesi, F. Danesin, F. Rampazzo, F. Zanon, A. Tazzoli, M. Meneghini, and E. Zanoni, "Reliability of GaN high-electron-mobility transistors:

state of the art and perspectives," *IEEE Transactions on Electron Devices*, vol. 8, no. 2, pp. 332–343, 2008.

[22] S. Binari, P. Klein, and T. Kazior, "Trapping effects in gan and sic microwave fets," *Proceedings of the IEEE*, vol. 90, no. 6, pp. 1048 – 1058, jun 2002.

[23] R. Vetury, N. Zhang, S. Keller, and U. Mishra, "The impact of surface states on the dc and rf characteristics of algan/gan hfets," *Electron Devices, IEEE Transactions on*, vol. 48, no. 3, pp. 560 –566, mar 2001.

[24] M. Khan, M. Shur, Q. Chen, and J. Kuznia, "Current/voltage characteristic collapse in algan/gan heterostructure insulated gate field effect transistors at high drain bias," *Electronics Letters*, vol. 30, no. 25, pp. 2175 –2176, dec 1994.

[25] J. Joh and J. a. del Alamo, "A Current-Transient Methodology for Trap Analysis for GaN High Electron Mobility Transistors," *IEEE Transactions on Electron Devices*, vol. 58, no. 1, pp. 132–140, Jan. 2011.

[26] J. Joh and J. A. del Alamo, "Mechanisms for electrical degradation of gan high-electron mobility transistors," in *Electron Devices Meeting, 2006. IEDM '06. International*, dec. 2006, pp. 1 –4.

[27] C. S. Smith, "Piezoresistance effect in germanium and silicon," *Phys. Rev.*, vol. 94, no. 1, pp. 42–49, Apr 1954.

[28] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Rev. Mod. Phys.*, vol. 54, no. 2, pp. 437–672, Apr 1982.

[29] M. V. Fischetti, F. Gámiz, and W. Hänsch, "On the enhanced electron mobility in strained-silicon inversion layers," *Journal of Applied Physics*, vol. 92, no. 12, pp. 7320–7324, 2002.

[30] G. Sun, *Ph.D. Dissertation, Strain effects on hole mobility of silicon and germanium p-type metal-oxide-semiconductor field-effect-transistors*, University of Florida, Gainesville, Fla., USA, 2007.

[31] K. Uchida, A. Kinoshita, and M. Saitoh, "Carrier transport in (110) nmosfets: Subband structures, non-parabolicity, mobility characteristics, and uniaxial stress engineering," *Electron Devices Meeting, 2006. IEDM '06. International*, pp. 1–3, Dec. 2006.

[32] S. L. R. Michael E. Levinshtein and M. S. Shur, *Properties of advanced semiconductor materials : GaN, AlN, InN, BN, SiC, SiGe*, Wiley, New York, NY, 2001.

[33] F. Balestra, S. Cristoloveanu, M. Benachir, J. Brini, and T. Elewa, "Double-gate silicon-on-insulator transistor with volume inversion: A new device with greatly enhanced performance," *Electron Device Letters, IEEE*, vol. 8, no. 9, pp. 410–412, Sep 1987.

[34] F. Stern and W. Howard, "Properties of semiconductor surface inversion layers in the electric quantum limit," *Physical Review*, 1967.

[35] S. Takagi, T. Mizuno, T. Tezuka, N. Sugiyama, S. Nakaharai, T. Numata, J. Koga, and K. Uchida, "Sub-band structure engineering for advanced cmos channels," *Solid-State Electronics*, vol. 49, no. 5, pp. 684 – 694, 2005, 5th International Workshop on the Ultimate Intergration of Silicon, ULIS 2004.

[36] S. Jin, M. V. Fischetti, and T. wei Tang, "Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band nonparabolicity," *Journal of Applied Physics*, vol. 102, no. 8, pp. 083715, 2007.

[37] T. Irisawa, T. Numata, T. Tezuka, K. Usuda, N. Sugiyama, and S.-I. , "Device design and electron transport properties of uniaxially strained-soi tri-gate nmosfets," *Electron Devices, IEEE Transactions on*, vol. 55, no. 2, pp. 649–654, Feb. 2008.

[38] A. Paul, M. Luisier, N. Neophytou, R. Kim, M. McLennan, M. Lundstrom, and G. Klimeck, "Band structure lab," May 2006.

[39] L. T. Canham, "Silicon quantum wire array fabrication by electrochemical and chemical dissolution of wafers," *Applied Physics Letters*, vol. 57, no. 10, pp. 1046–1048, 1990.

[40] A. R. Guichard, D. N. Barsic, S. Sharma, T. I. Kamins, and M. L. Brongersma, "Tunable light emission from quantum-confined excitons in tisi2-catalyzed silicon nanowires," *Nano Letters*, vol. 6, no. 9, pp. 2140–2144, 2006.

[41] T. Vo, A. J. Williamson, and G. Galli, "First principles simulations of the structural and electronic properties of silicon nanowires," *Physical Review B (Condensed Matter and Materials Physics)*, vol. 74, no. 4, pp. 045116, 2006.

[42] Y. Zheng, C. Rivas, R. Lake, K. Alam, T. Boykin, and G. Klimeck, "Electronic properties of silicon nanowires," *Electron Devices, IEEE Transactions on*, vol. 52, no. 6, pp. 1097–1103, June 2005.

[43] K. Majumdar and N. Bhat, "Bandstructure effects in ultra-thin-body double-gate field effect transistor: A fullband analysis," *Journal of Applied Physics*, vol. 103, no. 11, pp. 114503, 2008.

[44] T. Hara, Y. Yamada, T. Maegawa, and H. Tsuchiya, "Atomistic study on electronic properties of nanoscale soi channels," *Journal of Physics: Conference Series*, vol. 109, pp. 012012 (4pp), 2008.

[45] J. Wang, A. Rahman, A. Ghosh, G. Klimeck, and M. Lundstrom, "On the validity of the parabolic effective-mass approximation for the i-v calculation of silicon

nanowire transistors," *Electron Devices, IEEE Transactions on*, vol. 52, no. 7, pp. 1589–1595, July 2005.

[46] E. Gnani, S. Reggiani, A. Gnudi, P. Parruccini, R. Colle, M. Rudan, and G. Baccarani, "Band-structure effects in ultrascaled silicon nanowires," *Electron Devices, IEEE Transactions on*, vol. 54, no. 9, pp. 2243–2254, Sept. 2007.

[47] K. Nehari, N. Cavassilas, J. Autran, M. Bescond, D. Munteanu, and M. Lannoo, "Influence of band-structure on electron ballistic transport in silicon nanowire mosfet's: an atomistic study," *Solid-State Device Research Conference, 2005. ESSDERC 2005. Proceedings of 35th European*, pp. 229–232, 12-16 Sept. 2005.

[48] C. T. H. Scheel, S. Reich, "Electronic band structure of high-index silicon nanowires," *physica status solidi (b)*, vol. 242, no. 12, pp. 2474–2479, 2005.

[49] K.-H. Hong, J. Kim, S.-H. Lee, and J. K. Shin, "Strain-driven electronic band structure modulation of si nanowires," *Nano Letters*, vol. 8, no. 5, pp. 1335–1340, 2008.

[50] M.-F. Ng, L. Zhou, S.-W. Yang, L. Y. Sim, V. B. C. Tan, and P. Wu, "Theoretical investigation of silicon nanowires: Methodology, geometry, surface modification, and electrical conductivity using a multiscale approach," *Physical Review B (Condensed Matter and Materials Physics)*, vol. 76, no. 15, pp. 155435, 2007.

[51] N. Neophytou, A. Paul, M. Lundstrom, and G. Klimeck, "Bandstructure effects in silicon nanowire electron transport," *Electron Devices, IEEE Transactions on*, vol. 55, no. 6, pp. 1286–1297, June 2008.

[52] J.-A. Yan, L. Yang, and M. Y. Chou, "Size and orientation dependence in the electronic properties of silicon nanowires," *Physical Review B*, vol. 76, no. 11, pp. 115319–+, Sept. 2007.

[53] M. S. L. Neophytos Neophytou, Abhijeet Paul and G. Klimeck, "Simulations of nanowire transistors: atomistic vs. effective mass models," *Journal of Computational Electronics*, vol. 7, no. 3, pp. 363–366, 2008.

[54] C. Delerue, G. Allan, and M. Lannoo, "Theoretical aspects of the luminescence of porous silicon," *Phys. Rev. B*, vol. 48, no. 15, pp. 11024–11036, Oct 1993.

[55] G. Ledoux, O. Guillois, D. Porterat, C. Reynaud, F. Huisken, B. Kohn, and V. Paillard, "Photoluminescence properties of silicon nanocrystals as a function of their size," *Phys. Rev. B*, vol. 62, no. 23, pp. 15942–15951, Dec 2000.

[56] Y. Kanemitsu, "Efficient light emission from crystalline and amorphous silicon nanostructures," *Journal of Luminescence*, vol. 100, no. 1-4, pp. 209 – 217, 2002.

[57] N. Neophytou, A. Paul, and G. Klimeck, "Bandstructure effects in silicon nanowire hole transport," *Nanotechnology, IEEE Transactions on*, vol. 7, no. 6, pp. 710–719, Nov. 2008.

[58] P. W. Leu, A. Svizhenko, and K. Cho, "Ab initio calculations of the mechanical and electronic properties of strained si nanowires," *Physical Review B (Condensed Matter and Materials Physics)*, vol. 77, no. 23, pp. 235305, 2008.

[59] M. A. Reshchikov and H. Morkoç, "Luminescence properties of defects in gan," *Journal of Applied Physics*, vol. 97, no. 6, pp. 061301, 2005.

[60] C. G. Van de Walle, "First-principles calculations for defects and impurities: Applications to III-nitrides," *Journal of Applied Physics*, vol. 95, no. 8, pp. 3851, 2004.

[61] J. Neugebauer and C. G. V. de Walle, "Gallium vacancies and the yellow luminescence in gan," *Applied Physics Letters*, vol. 69, no. 4, pp. 503–505, 1996.

[62] C. Young, Y. Zhao, D. Heh, R. Choi, B. H. Lee, and G. Bersuker, "Pulsed id - vg methodology and its application to electron-trapping characterization and defect density profiling," *Electron Devices, IEEE Transactions on*, vol. 56, no. 6, pp. 1322 –1329, june 2009.

[63] H. N. Kim, *Ph.D. Dissertation, Qualitative and Quantative Characterization of Trapping Effects in AlGaN/GaN High Electron Mobility Transistors*, Ohio State University, Columbus, OH., USA, 2009.

[64] H. Kim and L. F. Eastman, "High-electric-field induced trap generation in algan/gan heterostructure field-effect transistors," *Journal of the Korean Physical Society*, vol. 55, no. 2, pp. 666–670, 2009.

[65] E. Nicollian and A. Goetzberger, "The Si-SiO2 interface-electrical properties as determined by the metal-insulator-silicon conductance technique," *Bell Syst. Tech. J*, vol. 46, no. 6, pp. 1055–1133, 1967.

[66] D. K. Schroder, *Semiconductor Material and Device Characterization, 3rd Edition*, Wiley, Hoboken, NJ, 2006.

[67] L. Semra, A. Telia, and A. Soltani, "Trap characterization in algan/gan hemt by analyzing frequency dispersion in capacitance and conductance," *Surface and Interface Analysis*, vol. 42, no. 6-7, pp. 799–802, 2010.

[68] E. J. Miller, X. Z. Dang, H. H. Wieder, P. M. Asbeck, E. T. Yu, G. J. Sullivan, and J. M. Redwing, "Trap characterization by gate-drain conductance and capacitance dispersion studies of an AlGaN/GaN heterostructure field-effect transistor," *Journal of Applied Physics*, vol. 87, no. 11, pp. 8070, 2000.

[69] R. M. Chu, Y. G. Zhou, K. J. Chen, and K. M. Lau, "Admittance characterization and analysis of trap states in AlGaN/GaN heterostructures," *Physica Status Solidi (C)*, vol. 2403, no. 7, pp. 2400–2403, Dec. 2003.

[70] E. Arslan, S. Bütün, Y. afak, and E. Ozbay, "Investigation of Trap States in AlInN/AlN/GaN Heterostructures by Frequency-Dependent Admittance Analysis," *Journal of Electronic Materials*, vol. 39, no. 12, pp. 2681–2686, Sept. 2010.

[71] J. Chung, X. Zhao, and T. Palacios, "Estimation of trap density in algan/gan hemts from subthreshold slope study," in *Device Research Conference, 2007 65th Annual*, june 2007, pp. 111 –112.

[72] C. Sah, L. Forbes, L. Rosier, and A. T. Jr., "Thermal and optical emission and capture rates and cross sections of electrons and holes at imperfection centers in semiconductors from photo and dark junction current and capacitance experiments," *Solid-State Electronics*, vol. 13, no. 6, pp. 759 – 788, 1970.

[73] C.T. and Sah, "Bulk and interface imperfections in semiconductors," *Solid-State Electronics*, vol. 19, no. 12, pp. 975 – 990, 1976.

[74] C. Sah and J. Walker, "Thermally stimulated capacitance for shallow majority-carrier traps in the edge region of semiconductor junctions," *Applied Physics Letters*, vol. 22, no. 8, pp. 384–385, 1973.

[75] C. T. Sah, W. W. Chan, H. S. Fu, and J. W. Walker, "Thermally stimulated capacitance (tscap) in p-n junctions," *Applied Physics Letters*, vol. 20, no. 5, pp. 193–195, 1972.

[76] D. V. Lang, "Deep-level transient spectroscopy: A new method to characterize traps in semiconductors," *Journal of Applied Physics*, vol. 45, no. 7, pp. 3023–3032, 1974.

[77] A. Hierro, D. Kwon, S. A. Ringel, M. Hansen, J. S. Speck, U. K. Mishra, and S. P. DenBaars, "Optically and thermally detected deep levels in n-type schottky and p[sup +]-n gan diodes," *Applied Physics Letters*, vol. 76, no. 21, pp. 3064–3066, 2000.

[78] A. R. Arehart, A. Corrion, C. Poblenz, J. S. Speck, U. K. Mishra, and S. A. Ringel, "Deep level optical and thermal spectroscopy of traps in n-gan grown by ammonia molecular beam epitaxy," *Applied Physics Letters*, vol. 93, no. 11, pp. 112101, 2008.

[79] A. Arehart, A. Sasikumar, G. Via, B. Winningham, B. Poling, E. Heller, and S. Ringel, "Spatially-discriminating trap characterization methods for hemts and their application to rf-stressed algan/gan hemts," in *Electron Devices Meeting (IEDM), 2010 IEEE International*, dec. 2010, pp. 20.1.1 –20.1.4.

[80] M. Faqir, G. Verzellesi, F. Fantini, F. Danesin, F. Rampazzo, G. Meneghesso, E. Zanoni, A. Cavallini, A. Castaldini, N. Labat, A. Touboul, and C. Dua, "Characterization and analysis of trap-related effects in algangan hemts," *Microelectronics Reliability*, vol. 47, no. 911, pp. 1639 – 1642, 2007.

[81] T. Okino, M. Ochiai, Y. Ohno, S. Kishimoto, K. Maezawa, and T. Mizutani, "Drain current dlts of algan-gan mis-hemts," *Electron Device Letters, IEEE*, vol. 25, no. 8, pp. 523 – 525, aug. 2004.

[82] P. B. Klein, S. C. Binari, J. A. Freitas, and A. E. Wickenden, "Photoionization spectroscopy of traps in GaN metal-semiconductor field-effect transistors," *Journal of Applied Physics*, vol. 88, no. 5, 2000.

[83] P. B. Klein, "Photoionization spectroscopy in AlGaN/GaN high electron mobility transistors," *Journal of Applied Physics*, vol. 92, no. 9, pp. 5498, 2002.

[84] G. Meneghesso, A. Chini, E. Zanoni, M. Manfredi, M. Pavesi, B. Boudart, and C. Gaquiere, "Diagnosis of trapping phenomena in gan mesfets," in *Electron Devices Meeting, 2000. IEDM Technical Digest. International*, 2000, pp. 389 –392.

[85] M. Wolter, P. Javorka, M. Marso, A. Alam, R. Carius, A. Fox, M. Heuken, H. Luth, and P. Kordos, "Investigations on the influence of traps in algan/gan hemts," in *Electron Devices for Microwave and Optoelectronic Applications, 2001 International Symposium on*, 2001, pp. 149 –154.

[86] D. Cheney, E. Douglas, B. Gila, F. Ren, and S. Pearton, "Algan/gan hemt trap detection with optical pumping," in *2011 Reliability of Compound Semiconductors (ROCS) Workshop*, may. 2011, pp. 1 –4.

[87] H. C. Yang, T. Y. Lin, and Y. F. Chen, "Nature of the 2.8-ev photoluminescence band in si-doped gan," *Phys. Rev. B*, vol. 62, pp. 12593–12596, Nov 2000.

[88] N. Yamamoto, H. Itoh, V. Grillo, S. F. Chichibu, S. Keller, J. S. Speck, S. P. DenBaars, U. K. Mishra, S. Nakamura, and G. Salviati, "Cathodoluminescence characterization of dislocations in gallium nitride using a transmission electron microscope," *Journal of Applied Physics*, vol. 94, no. 7, pp. 4315–4319, 2003.

[89] C. Díaz-Guerra, J. Piqueras, O. Volciuc, V. Popa, and I. M. Tiginyanu, "Cathodoluminescence microscopy and spectroscopy of gan epilayers microstructured using surface charge lithography," *Journal of Applied Physics*, vol. 100, no. 2, pp. 023509, 2006.

[90] M. Albrecht, J. L. Weyher, B. Lucznik, I. Grzegory, and S. Porowski, "Nonradiative recombination at threading dislocations in n-type gan: Studied by cathodoluminescence and defect selective etching," *Applied Physics Letters*, vol. 92, no. 23, pp. 231909, 2008.

[91] Z. H. Wu, A. M. Fischer, F. A. Ponce, B. Bastek, J. Christen, T. Wernicke, M. Weyers, and M. Kneissl, "Structural and optical properties of nonpolar gan thin films," *Applied Physics Letters*, vol. 92, no. 17, pp. 171904, 2008.

[92] R. Liu, A. Bell, F. A. Ponce, C. Q. Chen, J. W. Yang, and M. A. Khan, "Luminescence from stacking faults in gallium nitride," *Applied Physics Letters*, vol. 86, no. 2, pp. 021908, 2005.

[93] E. Douglas, S. Pearton, F. Ren, B. Poling, E. Heller, and D. Via, "Degradation of algan/gan high electron mobility transistors from x-band operation," in *Compound Semiconductor Integrated Circuit Symposium (CSICS), 2011 IEEE*, oct. 2011, pp. 1 –4.

[94] M. Yang, V. Chan, K. Chan, L. Shi, D. Fried, J. Stathis, A. Chou, E. Gusev, J. Ott, L. Burns, M. Fischetti, and M. Ieong, "Hybrid-orientation technology (hot): opportunities and challenges," *Electron Devices, IEEE Transactions on*, vol. 53, no. 5, pp. 965 – 978, may 2006.

[95] C. Young, M. Baykan, A. Agarwal, H. Madan, K. Akarvardar, C. Hobbs, I. Ok, W. Taylor, C. Smith, M. Hussain, T. Nishida, S. Thompson, P. Majhi, P. Kirsch, S. Datta, and R. Jammy, "Critical discussion on (100) and (110) orientation dependent transport: nmos planar and finfet," in *Symposium on VLSI Technology*, Kyoto, Japan, 2011.

[96] C. Kang, R. Choi, S. Song, K. Choi, B. Ju, M. Hussain, B. Lee, G. Bersuker, C. Young, D. Heh, P. Kirsch, J. Barnet, J.-W. Yang, W. Xiong, H.-H. Tseng, and R. Jammy, "A novel electrode-induced strain engineering for high performance soi finfet utilizing si (1hannel for both n and pmosfets," in *Electron Devices Meeting, 2006. IEDM '06. International*, dec. 2006, pp. 1 –4.

[97] K. Akarvardar, C. D. Young, M. O. Baykan, I. Ok, T. Ngai, K.-W. Ang, M. P. Rodgers, S. Gausepohl, P. Majhi, C. Hobbs, P. D. Kirsch, and R. Jammy, "Impact of fin doping and gate stack on finfet (110) and (100) electron and hole mobilities," *Electron Device Letters, IEEE*, vol. 33, no. 3, pp. 351 –353, march 2012.

[98] K. Uchida, A. Kinoshita, and M. Saitoh, "Carrier transport in (110) nmosfets: Subband structures, non-parabolicity, mobility characteristics, and uniaxial stress engineering," in *Electron Devices Meeting, 2006. IEDM '06. International*, dec. 2006, pp. 1 –3.

[99] L. Silvestri, S. Reggiani, E. Gnani, A. Gnudi, and G. Baccarani, "A low-field mobility model for bulk, ultrathin body soi and double-gate n-mosfets with different surface and channel orientations part i: Fundamental principles," *Electron Devices, IEEE Transactions on*, vol. 57, no. 7, pp. 1567 –1574, july 2010.

[100] K. Shimizu, T. Saraya, and T. Hiramoto, "Superior $\langle 110 \rangle$-directed electron mobility to $\langle 100 \rangle$-directed electron mobility in ultrathin body (110) n-type metal–oxide–semiconductor field-effect transistors," *Japanese Journal of Applied Physics*, vol. 49, no. 5, pp. 051303, 2010.

[101] J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, "Empirical $spds*$ tight-binding calculation for cubic semiconductors: General method and material parameters," *Phys. Rev. B*, vol. 57, no. 11, pp. 6493–6507, Mar 1998.

[102] T. B. Boykin, G. Klimeck, and F. Oyafuso, "Valence band effective-mass expressions in the $sp3d5s*$ empirical tight-binding model applied to a si and ge parametrization," *Phys. Rev. B*, vol. 69, no. 11, pp. 115201, Mar 2004.

[103] S. Timoshenko, *Strength of Materials*, Robert E. Krieger Publishing Co., Huntington, USA, 1976, c1958.

[104] M. Chu, T. Nishida, X. Lv, N. Mohta, and S. E. Thompson, "Comparison between high-field piezoresistance coefficients of si metal-oxide-semiconductor field-effect transistors and bulk si under uniaxial and biaxial stress," *Journal of Applied Physics*, vol. 103, no. 11, pp. 113704, 2008.

[105] Sematech, "International technology roadmap for semiconductors," 2008, Sematech.

[106] M. O. Baykan, S. E. Thompson, and T. Nishida, "Strain effects on three-dimensional, two-dimensional, and one-dimensional silicon logic devices: Predicting the future of strained silicon," *Journal of Applied Physics*, vol. 108, no. 9, pp. 093716, 2010.

[107] P. Hashemi, L. Gomez, and J. L. Hoyt, "Gate-all-around n-mosfets with uniaxial tensile strain-induced performance enhancement scalable to sub-10-nm nanowire diameter," *Electron Device Letters, IEEE*, vol. 30, no. 4, pp. 401–403, April 2009.

[108] N. Neophytou, S. G. Kim, G. Klimeck, and H. Kosina, "On the bandstructure velocity and ballistic current of ultra-narrow silicon nanowire transistors as a function of cross section size, orientation, and bias," *Journal of Applied Physics*, vol. 107, no. 11, pp. 113701, 2010.

[109] Y. Lee, K. Kakushima, K. Shiraishi, K. Natori, and H. Iwai, "Size-dependent properties of ballistic silicon nanowire field effect transistors," *Journal of Applied Physics*, vol. 107, no. 11, pp. 113705, 2010.

[110] A. Abudukelimu, K. Kakushima, P. Ahmet, K. Tsutsui, A. Nishiyama, N. Sugii, K. Natori, T. Hattori, and H. Iwai, "Performance of silicon ballistic nanowire mosfet with diverse orientations and diameters," *ECS Transactions*, vol. 27, no. 1, pp. 1111–1116, 2010.

[111] L. Zhang, H. Lou, J. He, and M. Chan, "Uniaxial strain effects on electron ballistic transport in gate-all-around silicon nanowire mosfets," *Electron Devices, IEEE Transactions on*, vol. 58, no. 11, pp. 3829 –3836, nov. 2011.

[112] K. Natori, "Compact modeling of ballistic nanowire mosfets," *Electron Devices, IEEE Transactions on*, vol. 55, no. 11, pp. 2877 –2885, nov. 2008.

[113] K. Natori, "Compact modeling of quasi-ballistic silicon nanowire mosfets," *Electron Devices, IEEE Transactions on*, vol. 59, no. 1, pp. 79 –86, jan. 2012.

[114] N. Neophytou and G. Klimeck, "Design space for low sensitivity to size variations in [110] pmos nanowire devices: The implications of anisotropy in the quantization mass," *Nano Letters*, vol. 9, no. 2, pp. 623–630, 2009, PMID: 19140767.

[115] J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, "Empirical $spds*$ tight-binding calculation for cubic semiconductors: General method and material parameters," *Phys. Rev. B*, vol. 57, no. 11, pp. 6493–6507, Mar 1998.

[116] T. B. Boykin, G. Klimeck, and F. Oyafuso, "Valence band effective-mass expressions in the $sp3d5s*$ empirical tight-binding model applied to a si and ge parametrization," *Phys. Rev. B*, vol. 69, no. 11, pp. 115201, Mar 2004.

[117] T. B. Boykin, M. Luisier, M. Salmani-Jelodar, and G. Klimeck, "Strain-induced, off-diagonal, same-atom parameters in empirical tight-binding theory suitable for [110] uniaxial strain applied to a silicon parametrization," *Phys. Rev. B*, vol. 81, pp. 125202, Mar 2010.

[118] S. Lee, F. Oyafuso, P. von Allmen, and G. Klimeck, "Boundary conditions for the electronic structure of finite-extent embedded semiconductor nanostructures," *Phys. Rev. B*, vol. 69, no. 4, pp. 045316, Jan 2004.

[119] A. Fischer and H. Richter, "BIAXIAL STATE OF STRESS IN SiGe EPILAYERS ON PATTERNED Si SUBSTRATES," *Materials Research*, vol. 318, pp. 703–708, 1994.

[120] J. J. Wortman and R. A. Evans, "Young's modulus, shear modulus, and poisson's ratio in silicon and germanium," *Journal of Applied Physics*, vol. 36, no. 1, pp. 153–156, 1965.

[121] J. C. Slater and G. F. Koster, "Simplified lcao method for the periodic potential problem," *Phys. Rev.*, vol. 94, no. 6, pp. 1498–1524, Jun 1954.

[122] T. B. Boykin, G. Klimeck, R. C. Bowen, and F. Oyafuso, "Diagonal parameter shifts due to nearest-neighbor displacements in empirical tight-binding theory," *Phys. Rev. B*, vol. 66, no. 12, pp. 125207, Sep 2002.

[123] T. B. Boykin, N. Kharche, and G. Klimeck, "Brillouin-zone unfolding of perfect supercells having nonequivalent primitive cells illustrated with a si/ge tight-binding parameterization," *Physical Review B (Condensed Matter and Materials Physics)*, vol. 76, no. 3, pp. 035310, 2007.

[124] I. Tiendaluna, F. Garciaruiz, L. Donetti, a. Godoy, and F. Gamiz, "Modeling the equivalent oxide thickness of Surrounding Gate SOI devices with high-κ insulators," *Solid-State Electronics*, vol. 52, no. 12, pp. 1854–1860, Dec. 2008.

[125] M. S. Lundstrom and J. Guo, *Nanoscale Transistors: Device Physics, Modeling and Simulation*, Springer, Norwood, MA, USA, 2006.

[126] M. Fischetti, S. Jin, T.-W. Tang, P. Asbeck, Y. Taur, S. Laux, M. Rodwell, and N. Sano, "Scaling mosfets to 10 nm: Coulomb effects, source starvation, and virtual source model," *Journal of Computational Electronics*, vol. 8, pp. 60–77, 2009, 10.1007/s10825-009-0277-z.

[127] J.-S. Lim, S. Thompson, and J. Fossum, "Comparison of threshold-voltage shifts for uniaxial and biaxial tensile-stressed n-mosfets," *Electron Device Letters, IEEE*, vol. 25, no. 11, pp. 731 – 733, nov. 2004.

[128] T. Ohta, H. Watanabe and I. Ohdomari, "Strain distribution around sio2/si interface in si nanowires: A molecular dynamics study," *Jpn. J. Appl. Phys.*, vol. 46, pp. 3277–3282, 2007.

[129] H. Minari, T. Zushi, T. Watanabe, Y. Kamakura, S. Uno, and N. Mori, "Impact of oxidation induced atomic disorder in narrow si nanowires on transistor performance," in *VLSI Technology (VLSIT), 2011 Symposium on*, june 2011, pp. 122 –123.

[130] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, "Surface roughness at the si(100)-$sio2$ interface," *Phys. Rev. B*, vol. 32, no. 12, pp. 8171–8186, Dec 1985.

[131] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the universality of inversion layer mobility in si mosfet's: Part i-effects of substrate impurity concentration," *Electron Devices, IEEE Transactions on*, vol. 41, no. 12, pp. 2357 –2362, dec 1994.

[132] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the universality of inversion layer mobility in si mosfet's: Part ii-effects of surface orientation," *Electron Devices, IEEE Transactions on*, vol. 41, no. 12, pp. 2363 –2368, dec 1994.

[133] X. Yang, S. Parthasarathy, Y. Sun, A. Koehler, T. Nishida, and S. E. Thompson, "Temperature dependence of enhanced hole mobility in uniaxial strained p-channel metal-oxide-semiconductor field-effect transistors and insight into the physical mechanisms," *Applied Physics Letters*, vol. 93, no. 24, pp. 243503, 2008.

[134] W. Chen, J. Kuo, and P. Su, "Experimental investigation of surface-roughness-limited mobility in uniaxial strained pmosfets," *Electron Device Letters, IEEE*, vol. 32, no. 2, pp. 113 –115, feb. 2011.

[135] O. Bonno, S. Barraud, D. Mariolle, and F. Andrieu, "Effect of strain on the electron effective mobility in biaxially strained silicon inversion layers: An experimental and theoretical analysis via atomic force microscopy measurements and

kubo-greenwood mobility calculations," *Journal of Applied Physics*, vol. 103, no. 6, pp. 063715, 2008.

[136] Y. Zhao, H. Matsumoto, T. Sato, S. Koyama, M. Takenaka, and S. Takagi, "A novel characterization scheme of interface roughness for surface roughness scattering-limited mobilities of electrons and holes in unstrained- and strained-si mosfets," *Electron Devices, IEEE Transactions on*, vol. 57, no. 9, pp. 2057 –2066, sept. 2010.

[137] Y. Zhao, M. Takenaka, and S. Takagi, "On surface roughness scattering-limited mobilities of electrons and holes in biaxially tensile-strained si mosfets," *Electron Device Letters, IEEE*, vol. 30, no. 9, pp. 987 –989, sept. 2009.

[138] U. Roh, T. Nishida, and S. E. Thompson, "Effect of applied strain on (100) silicon surface roughness: Physical insight into the enhanced electron mobility in strained nmosfets," *unpublished*.

[139] M. Shur, *Physics of semiconductor devices*, Prentice Hall, Englewood Cliffs, N.J, USA, 1990.

[140] R. Kotlyar, C. Weber, L. Shifren, S. Cea, M. D. Giles, and M. Stettler, "Effect of band warping and wafer orientation on nmos mobility under arbitrary applied stress," *Journal of Computational Electronics*, December 2007.

[141] J. Bardeen and W. Shockley, "Deformation potentials and mobilities in non-polar crystals," *Phys. Rev.*, vol. 80, no. 1, pp. 72–80, Oct 1950.

[142] G. L. Bir and G. E. Pikus, *Symmetry and strain-induced effects in semiconductors*, Wiley, New York, 1974.

[143] S. E. T. Yongke Sun and T. Nishida, *Strain Effects in Semiconductors: Theory and Device Applications*, Springer, London, UK, 2010.

[144] J. C. Hensel, H. Hasegawa, and M. Nakayama, "Cyclotron resonance in uniaxially stressed silicon. ii. nature of the covalent bond," *Phys. Rev.*, vol. 138, pp. A225–A238, Apr 1965.

[145] R. Kubo, "Statistical-mechanical theory of irreversible processes. i. general theory and simple applications to magnetic and conduction problems," *Journal of the Physical Society of Japan*, vol. 12, no. 6, pp. 570–586, 1957.

[146] D. A. Greenwood, "The boltzmann equation in the theory of electrical conduction in metals," *Proceedings of the Physical Society*, vol. 71, no. 4, pp. 585, 1958.

[147] 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*, vol. 55, pp. 645–705, 1983.

[148] S. M. G. David K. Ferry and J. Bird, *Transport in Nanostructures*, Cambridge University Press, Cambridge, UK, 2009.

[149] M. Fischetti and S. Laux, "Monte Carlo study of electron transport in silicon inversion layers," *Physical Review B*, vol. 48, no. 4, pp. 2244, 1993.

[150] D. Esseni, A. Abramo, L. Selmi, and E. Sangiorgi, "Physically based modeling of low field electron mobility in ultrathin single-and double-gate SOI n-MOSFETs," *Electron Devices, IEEE Transactions on*, vol. 50, no. 12, pp. 2445–2455, dec 2003.

[151] S. Yamakawa, H. Ueno, K. Taniguchi, C. Hamaguchi, K. Miyatsuji, K. Masaki, and U. Ravaioli, "Study of interface roughness dependence of electron mobility in si inversion layers using the monte carlo method," *Journal of Applied Physics*, vol. 79, no. 2, pp. 911–916, 1996.

[152] F. Gámiz, J. B. Roldán, J. A. López-Villanueva, P. Cartujo-Cassinello, and J. E. Carceller, "Surface roughness at the si–sio[sub 2] interfaces in fully depleted silicon-on-insulator inversion layers," *Journal of Applied Physics*, vol. 86, no. 12, pp. 6854–6863, 1999.

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

[154] F. Gámiz, J. B. Roldán, P. Cartujo-Cassinello, J. A. López-Villanueva, and P. Cartujo, "Role of surface-roughness scattering in double gate silicon-on-insulator inversion layers," *Journal of Applied Physics*, vol. 89, no. 3, pp. 1764–1770, 2001.

[155] S. Takagi, J. L. Hoyt, J. J. Welser, and J. F. Gibbons, "Comparative study of phonon-limited mobility of two-dimensional electrons in strained and unstrained si metal–oxide–semiconductor field-effect transistors," *Journal of Applied Physics*, vol. 80, no. 3, pp. 1567–1577, 1996.

[156] Y. Taur and T. H. Ning, *Fundamentals of modern VLSI devices*, Cambridge University Press, New York, NY, USA, 1998.

[157] E. Escobedo-Cousin, S. H. Olsen, T. Pardoen, U. Bhaskar, and J.-P. Raskin, "Experimental observations of surface roughness in uniaxially loaded strained si microelectromechanical systems-based structures," *Applied Physics Letters*, vol. 99, no. 24, pp. 241906, 2011.

[158] G. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, no. 8, pp. 114, Apr 1965.

[159] G. Moore, "Progress in digital integrated electronics," in *Electron Devices Meeting, 1975 International*, 1975, vol. 21, pp. 11–13.

[160] I. Corporation, "Intel 22nm 3-d tri-gate transistor technology," Tech. Rep., Intel Corporation, 2011.

[161] V. Zhirnov, I. Cavin, R.K., J. Hutchby, and G. Bourianoff, "Limits to binary logic switch scaling - a gedanken model," *Proceedings of the IEEE*, vol. 91, no. 11, pp. 1934–1939, Nov 2003.

[162] 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*, vol. 34, no. 4, pp. 466 –494, july 1990.

[163] M. Fischetti and S. Laux, "Monte carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures. ii. submicrometer mosfet's," *Electron Devices, IEEE Transactions on*, vol. 38, no. 3, pp. 650 –660, mar 1991.

[164] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, Wiley Interscience, 3 edition, January 2006.

[165] D. Chang and J. G. Fossum, "Simplified energy-balance model for pragmatic multi-dimensional device simulation," *Solid-State Electronics*, vol. 41, no. 11, pp. 1795–1802, 1997.

[166] C. Jacoboni and L. Reggiani, "The monte carlo method for the solution of charge transport in semiconductors with applications to covalent materials," *Rev. Mod. Phys.*, vol. 55, pp. 645–705, Jul 1983.

[167] M. V. Fischetti and S. E. Laux, "Monte carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Phys. Rev. B*, vol. 38, pp. 9721–9745, Nov 1988.

[168] J. P. McKelvey, R. L. Longini, and T. P. Brody, "Alternative approach to the solution of added carrier transport problems in semiconductors," *Phys. Rev.*, vol. 123, no. 1, pp. 51–57, Jul 1961.

[169] J. P. McKelvey, "Analysis of semiconductor p-n junctions and junction devices by a flux method," *Journal of Applied Physics*, vol. 33, no. 3, pp. 985–991, 1962.

[170] W. Shockley, "Diffusion and drift of minority carriers in semiconductors for comparable capture and scattering mean free paths," *Phys. Rev.*, vol. 125, no. 5, pp. 1570–1576, Mar 1962.

[171] J. P. McKelvey and J. C. Balogh, "Flux methods for the analysis of transport problems in semiconductors in the presence of electric fields," *Phys. Rev.*, vol. 137, no. 5A, pp. A1555–A1561, Mar 1965.

[172] M. Lundstrom and Z. Ren, "Essential physics of carrier transport in nanoscale mosfets," *Electron Devices, IEEE Transactions on*, vol. 49, no. 1, pp. 133 –141, jan 2002.

[173] F. Assad, K. Banoo, and M. Lundstrom, "The drift-diffusion equation revisited," *Solid-State Electronics*, vol. 42, no. 3, pp. 283 – 295, 1998.

[174] G. Timp, J. Bude, K. Bourdelle, J. Garno, A. Ghetti, H. Gossmann, M. Green, G. Forsyth, Y. Kim, R. Kleiman, F. Klemens, A. Kornblit, C. Lochstampfor, W. Mansfield, S. Moccio, T. Sorsch, D. Tennant, W. Timp, and R. Tung, "The ballistic nano-transistor," in *Electron Devices Meeting, 1999. IEDM Technical Digest. International*, 1999, pp. 55 –58.

[175] P. Palestri, D. Esseni, S. Eminente, C. Fiegna, E. Sangiorgi, and L. Selmi, "Understanding quasi-ballistic transport in nano-mosfets: Part I-scattering in the channel and in the drain," *Electron Devices, IEEE Transactions on*, vol. 52, no. 12, pp. 2727 – 2735, dec. 2005.

[176] S. Eminente, D. Esseni, P. Palestri, C. Fiegna, L. Selmi, and E. Sangiorgi, "Understanding quasi-ballistic transport in nano-mosfets: part II-technology scaling along the ITRS," *Electron Devices, IEEE Transactions on*, vol. 52, no. 12, pp. 2736 – 2743, dec. 2005.

[177] P. Palestri, R. Clerc, D. Esseni, L. Lucci, and L. Selmi, "Multi-subband-monte-carlo investigation of the mean free path and of the kt layer in degenerated quasi ballistic nanomosfets," in *Electron Devices Meeting, 2006. IEDM '06. International*, dec. 2006, pp. 1 –4.

[178] R. Clerc, P. Palestri, and L. Selmi, "On the physical understanding of the kt-layer concept in quasi-ballistic regime of transport in nanoscale devices," *Electron Devices, IEEE Transactions on*, vol. 53, no. 7, pp. 1634 –1640, july 2006.

[179] M.-J. Chen, H.-T. Huang, K.-C. Huang, P.-N. Chen, C.-S. Chang, and C. Diaz, "Temperature dependent channel backscattering coefficients in nanoscale mosfets," in *Electron Devices Meeting, 2002. IEDM '02. Digest. International*, 2002, pp. 39 – 42.

[180] H.-N. Lin, H.-W. Chen, C.-H. Ko, C.-H. Ge, H.-C. Lin, T.-Y. Huang, and W.-C. Lee, "Channel backscattering characteristics of strained pmosfets with embedded sige source/drain," in *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International*, dec. 2005, pp. 141 –144.

[181] H.-N. Lin, H.-W. Chen, C.-H. Ko, C.-H. Ge, H.-C. Lin, T.-Y. Huang, and W.-C. Lee, "Characterizing the channel backscattering behavior in nanoscale strained complementary metal oxide semiconductor field-effect transistors," *Japanese Journal of Applied Physics*, vol. 45, no. 11, pp. 8611–8617, 2006.

[182] Y. Tsai, S. Chung, P. Liu, C. Tsai, Y. Lin, C. Tsai, G. Ma, S. Chien, and S. Sun, "The channel backscattering characteristics of sub-100nm cmos devices with different channel/substrate orientations," in *VLSI Technology, Systems and Applications, 2007. VLSI-TSA 2007. International Symposium on*, april 2007, pp. 1 –2.

[183] M.-J. Chen, S.-G. Yan, R.-T. Chen, C.-Y. Hsieh, P.-W. Huang, and H.-P. Chen, "Temperature-oriented experiment and simulation as corroborating evidence of mosfet backscattering theory," *Electron Device Letters, IEEE*, vol. 28, no. 2, pp. 177 –179, feb. 2007.

[184] M. H. Liao, C. W. Liu, L. Yeh, T.-L. Lee, and M.-S. Liang, "Gate width dependence on backscattering characteristics in the nanoscale strained complementary metal-oxide-semiconductor field-effect transistor," *Applied Physics Letters*, vol. 92, no. 6, pp. 063506 –063506–3, feb 2008.

[185] R. Clerc, P. Palestri, L. Selmi, and G. Ghibaudo, "Back-scattering in quasi ballistic nanomosfets: The role of non thermal carrier distributions," in *Ultimate Integration of Silicon, 2008. ULIS 2008. 9th International Conference on*, march 2008, pp. 125 –128.

[186] M.-J. Chen, L.-F. Lu, and C.-Y. Hsu, "On the mean free path for backscattering in layer of bulk nano-mosfets," *Electron Devices, IEEE Transactions on*, vol. 55, no. 12, pp. 3594 –3598, dec. 2008.

[187] M.-J. Chen and L.-F. Lu, "A parabolic potential barrier-oriented compact model for the layer's width in nano-mosfets," *Electron Devices, IEEE Transactions on*, vol. 55, no. 5, pp. 1265 –1268, may 2008.

[188] V. Barral, T. Poiroux, J. Saint-Martin, D. Munteanu, J.-L. Autran, and S. Deleonibus, "Experimental investigation on the quasi-ballistic transport: Part I:determination of a new backscattering coefficient extraction methodology," *Electron Devices, IEEE Transactions on*, vol. 56, no. 3, pp. 408 –419, march 2009.

[189] V. Barral, T. Poiroux, D. Munteanu, J.-L. Autran, and S. Deleonibus, "Experimental investigation on the quasi-ballistic transport: Part II:backscattering coefficient extraction and link with the mobility," *Electron Devices, IEEE Transactions on*, vol. 56, no. 3, pp. 420 –430, march 2009.

[190] M. Vaidyanathan and D. L. Pulfrey, "An appraisal of the one-flux method for treating carrier transport in modern semiconductor devices," *Solid-State Electronics*, vol. 39, no. 6, pp. 827 – 832, 1996.

[191] M. Fischetti, S. Jin, T.-W. Tang, P. Asbeck, Y. Taur, S. Laux, M. Rodwell, and N. Sano, "Scaling mosfets to 10nm: Coulomb effects, source starvation, and virtual source model," *Journal of Computational Electronics*, vol. 8, pp. 60–77, 2009.

[192] K. Natori, "New solution to high-field transport in semiconductors: I. elastic scattering without energy relaxation," *Japanese Journal of Applied Physics*, vol. 48, no. 3, pp. 034503, 2009.

[193] K. Natori, "New solution to high-field transport in semiconductors: Ii. velocity saturation and ballistic transmission," *Japanese Journal of Applied Physics*, vol. 48, no. 3, pp. 034504, 2009.

[194] S. Thompson, M. Armstrong, C. Auth, M. Alavi, M. Buehler, R. Chau, S. Cea, T. Ghani, G. Glass, T. Hoffman, C.-H. Jan, C. Kenyon, J. Klaus, K. Kuhn, Z. Ma, B. Mcintyre, K. Mistry, A. Murthy, B. Obradovic, R. Nagisetty, P. Nguyen, S. Sivakumar, R. Shaheed, L. Shifren, B. Tufts, S. Tyagi, M. Bohr, and Y. El-Mansy, "A 90-nm logic technology featuring strained-silicon," *Electron Devices, IEEE Transactions on*, vol. 51, no. 11, pp. 1790 – 1797, nov. 2004.

[195] E. Karl, Y. Wang, Y.-G. Ng, Z. Guo, F. Hamzaoglu, U. Bhattacharya, K. Zhang, K. Mistry, and M. Bohr, "A 4.6ghz 162mb sram design in 22nm tri-gate cmos technology with integrated active vmin-enhancing assist circuitry," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, feb. 2012, pp. 230 –232.

[196] K. J. Kuhn, "Considerations for ultimate cmos scaling," *Electron Devices, IEEE Transactions on*, vol. PP, no. 99, pp. 1 –16, 2012.

[197] "The international technology roadmap for semiconductors," 2011.

[198] V. Barral, T. Poiroux, F. Rochette, M. Vinet, S. Barraud, O. Faynot, L. Tosti, F. Andrieu, M. Casse, B. Previtali, R. Ritzenthaler, P. Grosgeorges, E. Bernard, G. LeCarval, D. Munteanu, J. Autran, and S. Deleonibus, "Will strain be useful for 10nm quasi-ballistic fdsoi devices? an experimental study," in *VLSI Technology, 2007 IEEE Symposium on*, june 2007, pp. 128 –129.

[199] M. Saitoh, N. Yasutake, Y. Nakabayashi, K. Uchida, and T. Numata, "Understanding of strain effects on high-field carrier velocity in (100) and (110) cmosfets under quasi-ballistic transport," in *Electron Devices Meeting (IEDM), 2009 IEEE International*, dec. 2009, pp. 1 –4.

[200] E. Hsieh, S. Chung, P. Liu, W. Chiang, C. Tsai, W. Teng, C. Li, T. Kuo, Y. Wang, C. Yang, C. Tsai, and G. Ma, "A new and simple experimental approach to characterizing the carrier transport and reliability of strained cmos devices in the quasi-ballistic regime," in *Electron Devices Meeting (IEDM), 2009 IEEE International*, dec. 2009, pp. 1 –4.

[201] D. Fleury, G. Bidal, A. Cros, F. Boeuf, T. Skotnicki, and G. Ghibaudo, "New experimental insight into ballisticity of transport in strained bulk mosfets," in *VLSI Technology, 2009 Symposium on*, june 2009, pp. 16 –17.

[202] K. Natori, "Ballistic metal-oxide-semiconductor field effect transistor," *Journal of Applied Physics*, vol. 76, no. 8, pp. 4879–4890, 1994.

[203] J.-H. Rhew and M. S. Lundstrom, "Drift-diffusion equation for ballistic transport in nanoscale metal-oxide-semiconductor field effect transistors," *Journal of Applied Physics*, vol. 92, no. 9, pp. 5196 –5202, nov 2002.

[204] M. Shur, "Low ballistic mobility in submicron hemts," *Electron Device Letters, IEEE*, vol. 23, no. 9, pp. 511 – 513, sep 2002.

[205] M. Lundstrom, "Elementary scattering theory of the si mosfet," *Electron Device Letters, IEEE*, vol. 18, no. 7, pp. 361 –363, jul 1997.

[206] P. J. Price, "Chapter 4 monte carlo calculation of electron transport in solids," in *Lasers, Junctions, Transport*, R. Willardson and A. C. Beer, Eds., vol. 14 of *Semiconductors and Semimetals*, pp. 249 – 308. Elsevier, 1979.

[207] A. Rahman and M. Lundstrom, "A compact scattering model for the nanoscale double-gate mosfet," *Electron Devices, IEEE Transactions on*, vol. 49, no. 3, pp. 481 –489, mar 2002.

[208] M.-J. Chen, H.-T. Huang, Y.-C. Chou, R.-T. Chen, Y.-T. Tseng, P.-N. Chen, and C. Diaz, "Separation of channel backscattering coefficients in nanoscale mosfets," *Electron Devices, IEEE Transactions on*, vol. 51, no. 9, pp. 1409 – 1415, sept. 2004.

[209] H.-N. Lin, H.-W. Chen, C.-H. Ko, C.-H. Ge, H.-C. Lin, T.-Y. Huang, W.-C. Lee, and D. Tang, "The impact of uniaxial strain engineering on channel backscattering in nanoscale mosfets," in *VLSI Technology, 2005. Digest of Technical Papers. 2005 Symposium on*, june 2005, pp. 174 – 175.

[210] H.-W. Chen, H.-N. Lin, C.-H. Ko, C.-H. Ge, H.-C. Lin, T.-Y. Huang, and W.-C. Lee, "Strain-induced channel backscattering modulation in nanoscale cmosfets," in *VLSI Technology, Systems, and Applications, 2006 International Symposium on*, april 2006, pp. 1 –2.

[211] W. Lee and P. Su, "On the experimental determination of channel backscattering characteristics: Limitation and application for the process monitoring purpose," *Electron Devices, IEEE Transactions on*, vol. 56, no. 10, pp. 2285 –2290, oct. 2009.

[212] W. Lee, J.-Y. Kuo, W.-N. Chen, P. Su, and M.-C. Jeng, "Impact of uniaxial strain on channel backscattering characteristics and drain current variation for nanoscale pmosfets," in *VLSI Technology, 2009 Symposium on*, june 2009, pp. 112 –113.

[213] K. Natori, "Scaling limit of the MOS transistor - a ballistic MOSFET," *IEICE Transactions on Electronics*, vol. E84-C, no. 8, pp. 1029–1036, aug 2001.

[214] M. Fischetti, S. Jin, T.-w. Tang, P. Asbeck, Y. Taur, S. Laux, and N. Sano, "Scaling mosfets to 10 nm: Coulomb effects, source starvation, and virtual source," in *Computational Electronics, 2009. IWCE '09. 13th International Workshop on*, may 2009, pp. 1 –4.

[215] S. Datta, *Electronic Transport in Mesoscopic Systems*, Cambridge University Press, 1 edition, May 1997.

[216] F. Berz, "The bethe condition for thermionic emission near an absorbing boundary," *Solid-State Electronics*, vol. 28, no. 10, pp. 1007 – 1013, 1985.

[217] R. Kim and M. Lundstrom, "Physics of carrier backscattering in one- and two-dimensional nanotransistors," *Electron Devices, IEEE Transactions on*, vol. 56, no. 1, pp. 132 –139, jan. 2009.

[218] A. Svizhenko and M. Anantram, "Role of scattering in nanotransistors," *Electron Devices, IEEE Transactions on*, vol. 50, no. 6, pp. 1459 – 1466, june 2003.

[219] G. Gildenblat, "One-flux theory of a nonabsorbing barrier," *Journal of Applied Physics*, vol. 91, no. 12, pp. 9883–9886, 2002.

[220] E. F. Pulver and J. P. McKelvey, "Flux methods for transport problems in solids with nonconstant electric fields," *Phys. Rev.*, vol. 149, no. 2, pp. 617–623, Sep 1966.

[221] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Rev. Mod. Phys.*, vol. 54, pp. 437–672, Apr 1982.

[222] D. Esseni, P. Palestri, and L. Selmi, *Nanoscale MOS Transistors: Semi-Classical Transport and Applications*, Cambridge University Press, 1 edition, Jan 2011.

[223] S. H. Wemple and J. A. Seman, "Optical transmission through multilayered structures," *Appl. Opt.*, vol. 12, no. 12, pp. 2947–2949, Dec 1973.

[224] I. Nagrath and M. Gopal, *Control System Engineering*, Wiley Eastern Limited, 2 edition, May 1993.

[225] E. J. Ryder, "Mobility of holes and electrons in high electric fields," *Phys. Rev.*, vol. 90, no. 5, pp. 766–769, Jun 1953.

[226] H. Wang and G. Gildenblat, "Scattering matrix based compact mosfet model," in *Electron Devices Meeting, 2002. IEDM '02. Digest. International*, 2002, pp. 125 – 128.

[227] G. Gildenblat, X. Li, W. Wu, H. Wang, A. Jha, R. van Langevelde, G. Smit, J. Scholten, and D. Klaassen, "Psp: An advanced surface-potential-based mosfet model for circuit simulation," *Electron Devices, IEEE Transactions on*, vol. 53, no. 9, pp. 1979 –1993, sept. 2006.

[228] G. Gildenblat, *Compact Modeling: Principles, Techniques and Applications*, Springer, 2010.

[229] Y. Sun, S. E. Thompson, and T. Nishida, *Strain Effects in Semiconductors: Theory and Device Applications*, Springer, 2010.

[230] M. O. Baykan, S. E. Thompson, and T. Nishida, "Strain effects on three-dimensional, two-dimensional, and one-dimensional silicon logic devices: Predicting the future of strained silicon," *Journal of Applied Physics*, vol. 108, no. 9, pp. 093716, 2010.

[231] Y. Sun, G. Sun, S. Parthasarathy, and S. Thompson, "Physics of process induced uniaxially strained si," *Materials Science and Engineering: B*, vol. 135, no. 3, pp. 179 – 183, 2006.

[232] S. Thompson, G. Sun, Y. S. Choi, and T. Nishida, "Uniaxial-process-induced strained-si: extending the cmos roadmap," *Electron Devices, IEEE Transactions on*, vol. 53, no. 5, pp. 1010–1020, May 2006.

[233] G. Sun, Y. Sun, T. Nishida, and S. E. Thompson, "Hole mobility in silicon inversion layers: Stress and surface orientation," *Journal of Applied Physics*, vol. 102, no. 8, pp. 084501, 2007.

[234] S. Flachowsky, A. Wei, R. Illgen, T. Herrmann, J. Ho andntschel, M. Horstmann, W. Klix, and R. Stenzel, "Understanding strain-induced drive-current enhancement in strained-silicon n-mosfet and p-mosfet," *Electron Devices, IEEE Transactions on*, vol. 57, no. 6, pp. 1343 –1354, june 2010.

[235] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson, and M. Bohr, "A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon cmos transistors," in *Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International*, dec. 2003, pp. 11.6.1 – 11.6.3.

[236] Y.-F. Wu, B. Keller, S. Keller, D. Kapolnek, S. Denbaars, and U. Mishra, "Measured microwave power performance of algan/gan modfet," *Electron Device Letters, IEEE*, vol. 17, no. 9, pp. 455 –457, sep 1996.

[237] N. Faleev and I. Levin, "Strain and crystal defects in thin aln/gan structures on (0001) sic," *Journal of Applied Physics*, vol. 107, no. 11, pp. 113529, 2010.

[238] N. Faleev, C. Honsberg, O. Jani, and I. Ferguson, "Crystalline perfection of gan and aln epitaxial layers and the main features of structural transformation of crystalline defects," *Journal of Crystal Growth*, vol. 300, no. 1, pp. 246 – 250, 2007.

[239] S. Cho, K. Chang, and M. Kwon, "Strain analysis of a gan epilayer grown on a c-plane sapphire substrate with different growth times," *Journal of Materials Science*, vol. 42, pp. 3569–3572, 2007, 10.1007/s10853-007-1562-5.

[240] C. Kisielowski, J. Krüger, S. Ruvimov, T. Suski, J. W. Ager, E. Jones, Z. Liliental-Weber, M. Rubin, E. R. Weber, M. D. Bremser, and R. F. Davis, "Strain-related phenomena in gan thin films," *Phys. Rev. B*, vol. 54, pp. 17745–17753, Dec 1996.

[241] T. Y. Lin, H. C. Yang, and Y. F. Chen, "Optical quenching of the photoconductivity in n-type gan," *Journal of Applied Physics*, vol. 87, no. 7, pp. 3404–3408, 2000.

[242] M. Wolter, P. Javorka, M. Marso, a. Alam, R. Carius, a. Fox, M. Heuken, H. Luth, and P. Kordos, "Investigations on the influence of traps in AlGaN/GaN HEMTs," *2001 International Symposium on Electron Devices for Microwave and Optoelectronic Applications. EDMO 2001 (Cat. No.01TH8567)*, pp. 149–154, 2001.

[243] C.-Y. Chang, E. Douglas, J. Kim, L. Lu, C.-F. Lo, B.-H. Chu, D. Cheney, B. Gila, F. Ren, G. Via, D. Cullen, L. Zhou, D. Smith, S. Jang, and S. Pearton, "Electric-field-driven degradation in off-state step-stressed algan/gan high-electron mobility transistors," *Device and Materials Reliability, IEEE Transactions on*, vol. 11, no. 1, pp. 187 –193, march 2011.

BIOGRAPHICAL SKETCH

Mehmet Onur Baykan was born in Adana, Turkey. He has received the B.S. degree from the Department of Electrical and Electronics Engineering at Middle East Technical University, Ankara/Turkey in 2007, and M.S and Ph.D. degrees from the Department of Electrical and Computer Engineering at University of Florida in 2009 and 2012 respectively. He was an intern Device Engineer at SEMATECH from June 2010 to May 2011. He is currently with Intel Portland Technology Development at Hillsboro, OR, working on 22nm, 14nm and 10nm technology node integration, development and yield enhancement. His research interests include carrier transport and reliability in ultra-scaled strained silicon nanodevices, non-planar silicon and compound semiconductor field-effect-transistors.