Monte Carlo Study of Current Variability in UTB SOI DG MOSFETs

Craig Riddet

Submitted to the Faculty of Engineering, Department of Electronics and Electrical Engineering, University of Glasgow in fulfillment of the requirements of the degree of Doctor of Philosophy.

January 2008

All work © Craig Riddet, 2008

Abstract

The scaling of conventional silicon based MOSFETs is increasingly difficult into the nanometer regime due to short channel effects, tunneling and subthreshold leakage current. Ultra-thin body silicon-on-insulator based architectures offer a promising alternative, alleviating these problems through their geometry. However, the transport behaviour in these devices is more complex, especially for silicon thicknesses below 10 nm, with enhancement from band splitting and volume inversion competing with scattering from phonons, Coulomb interactions, interface roughness and body thickness fluctuation.

Here, the effect of the last scattering mechanism on the drive current is examined as it is considered a significant limitation to device performance for body thicknesses below 5 nm. A simulation technique that properly captures non-equilibrium transport, includes quantum effects and maintains computational efficiency is essential for the study of this scattering mechanism. Therefore, a 3D Monte Carlo simulator has been developed which includes this scattering effect in an *ab initio* fashion, and quantum corrections using the Density Gradient formalism. Monte Carlo simulations using 'frozen field' approximation have been carried out to examine the dependence of mobility on silicon thickness in large, self averaging devices. This approximation is then used to carry out statistical studies of uniquely different devices to examine the variability of on-current. Finally, Monte Carlo simulations self consistent with Poisson's equation have been carried out to further investigate this mechanism.

Acknowledgements

This work would not have been possible without an assortment of rather wonderful people. Firstly, a great deal of thanks goes to my supervisors, Prof. Asen Asenov and Dr Scott Roy for their patience and guidance over the duration of my studies. And of course, a great deal of thanks goes to the various members of the Device Modelling Group (too numerous to mention), who have been on hand to answer the many questions I've asked over the last few years. I'm grateful too to the support from Fujitsu which made this project possible in the first place. And of course, friends and family who have constantly and patiently listened to me claim that this thesis would one day be finally finished!

List of Publications

1. C. Riddet, A. R. Brown, C. Alexander, J. R. Watling, S. Roy and A. Asenov, "3-D Monte Carlo Simulation of the Impact of Quantum Confinement Scattering on the Magnitude of Current Fluctuations in Double Gate MOSFETs," *IEEE Transactions on Nanotechnology*, vol. 6, No. 1, pp. 48-55, 2007.

2. C. Riddet, A. R. Brown, C. Alexander, S. Roy and A. Asenov, "Efficient Density Gradient Quantum Corrections for 3D Monte Carlo Simulations," *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 200-203, 2006.

3. C. Riddet, A. R. Brown, C. Alexander, J. R. Watling, S. Roy and A. Asenov, "Impact of Quantum Confinement Scattering on the Magnitude of Current Fluctuations in Double Gate MOSFETs," *Silicon Nanoelectronics Workshop, 12-13 June, Kyoto, Japan*, pp. 6-7, 2005.

4. C. Riddet, A. Brown, C. Alexander, J. R. Watling, S. Roy and A. Asenov, "Scattering from Body Thickness Fluctuations in Double Gate MOSFETs. An *ab initio* Monte Carlo Simulation Study," *Journal of Computational Electronics*, Vol. 3, pp. 341-345, 2004.

5. C. Riddet, A. Brown, C. Alexander, J. R. Watling, S. Roy and A. Asenov, "Scattering from Body Thickness Fluctuations in Double Gate MOSFETs. An *ab initio* Monte Carlo Simulation Study," *Abstracts of the 10th International Workshop on Computational Electronics (IWCE-10), W. Lafayette, IN, USA, 24-27 October*, pp. 194-195, 2004.

Conference Papers

1. **12th International Workshop on Computational Electronics (IWCE 12) (poster presentation)**, University of Massachusetts, Amherst, USA, October 2007. "Boundary Conditions for Density Gradient Corrections in 3D Monte Carlo Simulations".

2. **International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) (poster presentation)**, Monterey, Califorina, USA, September 2006. "Efficient Density Gradient Quantum Corrections for 3D Monte Carlo Simulations".

3. **Silicon Nanoelectronics Workshop (oral presentation)**, Kyoto, Japan, June 2005. "Impact of Quantum Confinement Scattering on the Magnitude of Current Fluctuations in Double Gate MOSFETs".

4. **10th International Workshop on Computational Electronics (IWCE 10) (poster presentation)**, Purdue University, West Layafette, USA, October 2004. "Scattering from Body Thickness Fluctuations in Double Gate MOSFETs. An *ab initio* Monte Carlo Simulation Study".

Contents

1 Introduction **1**

1.1 Scaling of MOSFETs and the Transition to New Device Architectures 1

1.2 Aims and Objectives 4

1.3 Thesis Outline 4

2 Scaling and Transport Phenomena of UTB Devices **6**

2.1 Introduction 6

2.2 Scaling Properties of Conventional, UTB SOI and DG MOSFETs 7

2.3 Volume Inversion 12

2.4 Enhanced Phonon Scattering 14

2.5 Band Splitting 17

2.6 Coulomb Scattering 20

2.7 Interface Roughness and Body Thickness Variation 22

2.8 Summary 26

3 Simulation Methodology 29

3.1 Introduction 29

3.2 Drift-Diffusion 31

3.2.1 Quantum Corrections in Drift-Diffusion 34

3.3 Monte Carlo 38

3.3.1 Scattering in Monte Carlo 41

3.3.2 Quantum Corrections in Monte Carlo 43

3.4 Quantum Transport 49

3.4.1 Scattering in NEGF 52

3.5 Summary 54

4 Simulator Development 56

4.1 Introduction 56

4.2 Drift Diffusion 57

4.2.1 Interface Roughness Generation 60

4.3 Monte Carlo 62

4.3.1 'Frozen Field' Approximation 63

4.3.2 Interface Roughness Implementation 65

4.3.3 Self-Consistent Monte Carlo 66

4.3.4 Quantum Corrections 68
4.3.5 Ohmic Contacts 75
4.3.6 Additional Considerations 82
4.3.7 Summary 83

5 Results And Discussion 85

5.1 Introduction 85
5.2 Mobility Dependence on Silicon Thickness 86
5.3 Current Variation in DG and SOI MOSFETs 94
5.3.1 'Frozen Field' Monte Carlo Simulation Studies 94
5.3.2 Self-Consistent Monte Carlo Simulation Studies 100
5.4 Summary 106

6 Conclusions 108

6.1 Suggestions for Future Work 111

Bibliography 138

List of Figures

2.1 Schematics of SG-SOI and DG architectures. 9

2.2 Comparison of potentials and electron distributions in DG MOSFETs. 13

2.3 Phonon limited mobility versus t_{Si}. 15

2.4 Dependence of phonon limited electron mobility on silicon layer thickness. 16

2.5 Mobility in different subbands and overall mobility. 18

2.6 Subband structure, showing 4- (in plane) and 2- (out plane) fold valleys in red and blue respectively. 20

2.7 Comparison of 2- and 4-fold valleys showing energy separation. . 21

2.8 Source of body thickness fluctuations resulting from roughness patterns. 23

2.9 Nominal and roughness limited thickness of silicon layer. 24

2.10 Mobility dependence on t_{Si}. 25

2.11 Comparison of mobility vs. silicon thickness trends in SG-SOI MOSFETs. 26

2.12 Comparison of mobility vs. silicon thickness trends in DG MOSFETs. 27

3.1 Electron distribution from NEGF and Density Gradient simulations. 35

3.2 Simulation flow for Drift Diffusion using a modified Gummel algorithm 36

3.3 Simulation flow for Monte Carlo. 38

3.4 Relationship of scattering rates to random number. 40

3.5 Cumulative scattering rates in MC. 41

3.6 Simulation flow for NEGF. 51

3.7 Hierarchy of simulation methodologys. 54

4.1 Generated roughness pattern and the corresponding digitised version. 60

4.2 Confinement due to interface roughness. 61

4.3 Digitization of interface roughness patterns in MC. 65

4.4 Vertical quantum correction. 69

4.5 Flowchart of fully self-consistent MC simulation. 72

4.6 Timeline of fully self-consistent MC simulation. 73

4.7 Quantum corrected potential and electron concentration in a DG MOSFET. 74

4.8 Electron concentration profile from poorly maintained contacts. 76

4.9 Position and orientation of ohmic contacts for a MC simulation. 77

4.10 Comparison of characteristics with and without reservoir contacts. 78

4.11 The electron concentration at the source contact in the vertical direction. 79

4.12 $I_D - V_G$ for a single device using both methods of boundary condition. 80

4.13 Electron concentration profile from well maintained contacts. . . 81

4.14 Various methods for dealing with boundary conditions in Density Gradient. 82

5.1 Devices simulated in this study. 87

5.2 3D plot of electron distribution. 88

5.3 Classical and Quantum potential planes. 89

5.4 Potential profile through the centre of the device. 90

5.5 Potential plane running from the source to the drain. 90

5.6 Mobility dependence on silicon thickness in DG MOSFET. . . . 91

5.7 Electron distribution and quantum potential in $t_{Si} = 2.4$ nm device. 92

5.8 Comparison of simulated to theoretical dependence of mobility on t_{Si}. . . . 93

5.9 Characteristics from DD and FFMC with uniform body thickness. 95

5.10 $I_D - V_G$ characteristics for DG MOSFET comparing DD and FFMC. . . . 96

5.11 Comparison of I_D variation for 50 uniquely different DG MOSFETs. . . . 97

5.12 Comparison of I_D variation for 50 uniquely different SOI MOSFETs. . . . 98

5.13 I_D percentage variation for 50 uniquely different DG MOSFETs. 99

5.14 I_D percentage variation for 50 uniquely different SOI MOSFETs. 100

5.15 3D plot of classical potential. . . . 101

5.16 $I_D - V_G$ characteristics comparing DD, self consistent and FFMC. 101

5.17 Carrier concentration on a plane in the channel (uniform thickness). . . . 103

5.18 DD carrier concentration on a plane in the channel (non-uniform thickness). . . . 103

5.19 MC carrier concentration on a plane in the channel (non-uniform thickness). . . . 104

5.20 3D plot of quantum corrected potential, as calculated in the MC simulator, at $V_D = 1$ mV, $V_G = 0.8$V, with the units also in Volts. 105

5.21 3D plot of electron concentration. . . . 105

List of Tables

2.1 Comparison of DIBL in various MOSFET structures. 12

2.2 Description of the subband structure for (100) Si orientation. . . 19

3.1 Strengths and weaknesses of quantum correction methods in MC. 49

4.1 Comparison of interface roughness implementation in various studies. 59

4.2 Average time of solution for Poisson's equation with different solvers. 67

4.3 Comparison of the average CPU time to carry out a time step for each simulation method. 75

5.1 Dimensions of devices simulated in this study. 86

5.2 Statistics from the simulation of 50 DG devices. 98

5.3 Statistics from the simulation of 50 SOI devices. 99

5.4 Comparison of quantum corrected DD and MC simulations . . . 102

Glossary

Acronyms and Abbreviations

Bi-CGSTAB Bi-Conjugate Stabilised

BTV Body Thickness Variations

DD Drift Diffusion

DG Double Gate

DIBL Drain Induced Barrier Lowering

ECBE Effective Conduction Band Edge

FD Fully Depleted

FFMC 'Frozen Field' Monte Carlo

FQMC Frozen Quantum Correction Monte Carlo

MC Monte Carlo

MOSFET Metal Oxide Semiconductor Field Effect Transistor

NEGF Non Equilibrium Green's Function

PD Partially Depleted

SCE Short Channel Effects

SCQMC Self-Consistent Quantum Monte Carlo

SOI Silicon On Insulator

SOR Successive Over Relaxation

UTB Ultra Thin Body

List Of Symbols

E Energy [eV]

E_0 Ground State Energy [eV]

ε Material Permittivity [F/cm]

ϕ_n Quasi-Fermi Level [V]

I_D Drain Current [A/μm]

I_{off} Off Current [A/μm]

I_{on} On Current [A/μm]

L_g Gate Length [nm]

m^* Effective Mass [kg]

μ Carrier Mobility [cm^2/Vs]

μ_n Electron Mobility [cm^2/Vs]

n Electron Concentration [cm^{-3}]

N_A Acceptor Concentration [cm^{-3}]

N_D	Donor Concentration [cm^{-3}]
n	Electron Density [cm^{-3}]
n_{mc}	Electron Density in Monte Carlo [cm^{-3}]
n_q	Quantum Density [cm^{-3}]
p	Hole Concentration [cm^{-3}]
T	Temperature [K]
t_{ox}	Oxide Thickness [nm]
t_{Si}	Silicon Layer Thickness [nm]
$\mathbf{v}$	Carrier Velocity [cm^{2}]
V_D	Drain Voltage [V]
V_G	Gate Voltage [V]
V_T	Threshold Voltage [V]
ψ	Potential [V]
ψ_c	Classical Potential [V]
ψ_q	Quantum Potential [V]
ψ_{qc}	Quantum Correction Potential [V]
$\langle ... \rangle_t$	Time Averaged Value [-]

List Of Physical Constants

ε_{ox}	Permittivity of Oxide [= 3.45 $\times 10^{-13}$ F/cm]
ε_{Si}	Permittivity of Silicon [= 1.04 $\times 10^{-12}$ F/cm]
h	Plank's Constant [= 6.63 $\times 10^{-34}$ J-s]
$\hbar$	Reduced Plank's Constant [= 1.05 $\times 10^{-34}$ J-s]
k_B	Boltzmann's Constant [= 1.38 $\times 10^{-23}$ J/K]
m_0	Mass of an Electron [= 9.1 $\times 10^{-31}$ kg]
q	Electronic Charge [= 1.6 $\times 10^{-19}$ C]

Chapter 1

Introduction

The aim of this work is the development of a Monte Carlo simulator capable of studying the current variability in ultra thin body (UTB) silicon on insulator (SOI) and double gate (DG) MOSFETs which are considered strong candidates to replace the conventional MOSFET architecture as devices are scaled down to increasingly small dimensions.

In this section a brief overview of various possibilities for scaling of MOSFETs further into the nanometer regime, based on the adoption of alternative materials and device structures, is presented. Following this, the aims and objectives of this study are outlined in more detail, and an overview of the structure and content of this thesis is given.

1.1 Scaling of MOSFETs and the Transition to New Device Architectures

The scaling of conventional silicon (Si) based MOSFETs becomes increasingly difficult in the sub-100 nm channel length regime, as short channel effects be-

come ever more problematic and quantum mechanical tunneling through the gate oxide and band-to-band tunneling leads to greater gate and subthreshold leakage current. The desire to maintain good on-current (I_{on}), while keeping subthreshold currents (I_{off}) as small as possible has resulted in an increased interest in alternate materials (III-V, high-κ dielectrics, strained Si) and architectures (based on SOI and multiple gate designs) that are more resistant to short channel effects, and capable of maintaining high performance at small dimensions [1].

The use of Si under biaxial tensile strain on SiGe (silicon-germanium) substrates has been extensively investigated [2, 3, 4, 5, 6]. Strained Si produces a higher carrier mobility as a result of a splitting of the high and low effective mass valleys [7], hence enhanced transport compared to conventional Si MOSFETs. Difficulties in terms of developing fabrication processes that minimize defects [8], along with the observation of self-heating in such devices [9] are, however, significant drawbacks. Instead, process induced uniaxial strain is now widely used to enhance the device performance [10, 11].

Another possibility comes in replacing Si altogether. MOSFETs employing compound semiconductor (III-V) materials such as gallium arsenide (GaAs) [12], indium gallium arsenide (InGaAs) [13] or indium antimonide (InSb) [14] are attractive due to their higher saturation velocity and mobility. However, Si is cheaper and easier to process due to its greater strength which allows for larger wafers to be manufactured, and offers an excellent native oxide in SiO_2, factors that give it significant advantage in terms of the fabrication of devices, and that are still being addressed for III-V technology with the use of metal gates and high-κ dielectrics [15]. Therefore, III-V materials have a potential if only they can be integrated as channel materials on large Si substrates. Although III-V MOSFETs with conventional architecture offer performance improvements only down to 30 nm channel lengths [16, 17], new implant free architectures offer nanometer scaling potential [18, 19]. The smaller band gap and higher dielectric constant of these materials means they can also be more

susceptible to band-to-band tunneling and short channel effects [1, 17], both already considered a major limitations to the scaling of Si devices.

The inclusion of high-κ dielectrics [20] in both III-V and conventional Si MOSFETs has also gained a lot of interest. A reduction in gate leakage current comes via the oxide layer being physically thicker than conventional silicon dioxide (SiO_2) while maintaining a similar capacitance. However, finding a high-κ material with qualities comparable (in terms of potential barrier height, thermal stability and interface properties [7]) to SiO_2 is difficult and the use of such dielectrics adds an additional source of carrier scattering from soft optical phonons which results in degradation of carrier mobility [21, 22].

The use of alternative architectures [7, 23] is attractive as they generally remain based around Si and SiO_2 (although the use of III-V, strain and high-κ has been investigated in these designs [17, 24, 25]), well understood materials long used in conventional MOSFETs. A range of possible designs have been proposed from the single gated silicon-on-insulator (SOI) structure, to the double gate (DG) configuration, FinFET, triple-gate and gate-all-around (GAA) designs [23]. These alternate architectures allow for better suppression of short channel effects through the device geometry and negate issues associated with highly doped regions in conventional structures by using virtually undoped channels, allowing for higher carrier mobility and less variability.

In this thesis the single- and double-gated ultra thin body (UTB) configurations that are expected to replace conventional MOSFETs from the 32 nm technology node [26, 27], are investigated in detail. As these devices are scaled to silicon thicknesses below 10 nm, transport behaviour becomes increasingly complex, with enhancement from volume inversion and band splitting competing with enhanced scattering from Phonons, Coulomb interactions with charged centres trapped at the semiconductor/oxide interfaces and surface roughness and body thickness fluctuation induced scattering. All of these come about due to the design and structure of these devices, so a complete un-

derstanding of the full impact of these mechanisms is vital in order to develop an accurate picture of how well SOI MOSFETs will scale.

1.2 Aims and Objectives

As SOI and DG MOSFETs are scaled to nanometer scale Si body thicknesses, the influence of the roughness patterns present at the silicon/oxide interfaces because an increasingly significant issue. In addition to scattering directly from the roughness patterns, the non-uniform silicon thickness along the channel leads to a shift in the ground state and a resulting fluctuation in the effective quantum potential that acts as an additional source of carrier scattering, thus degrading transport [28]. This is one of the transport degrading phenomena considered fundamental to the limits of scaling of SOI and DG MOSFETs.

The aim of this project is to a develop computationally efficient methods to study this phenomena based on the Monte Carlo technique which is capable of capturing non equilibrium transport and scattering effects. In order to take into account the variable shifting of the ground state that results in additional scattering, the inclusion of quantum mechanical effects is of great importance. To this end, techniques have been developed and implemented based on the Density Gradient formalism. Using this simulation methodology, the impact of scattering from body thickness fluctuations on carrier transport and device variability has been examined.

1.3 Thesis Outline

This thesis is structured as follows. In Chapter 2 a more detailed examination of the scaling and transport behaviour associated with UTB SOI and DG MOSFETs is given. The scaling of conventional, SOI and DG architectures

are compared and contrasted, demonstrating the benefits in moving to the new structures as device dimensions shrink. After that, a review of previous studies on transport phenomena in UTB SOI and DG devices is presented, looking at the various mechanisms that dictate carrier transport in these architectures as the silicon layer thickness is scaled down.

Chapter 3 provides an evaluation of simulation methodologies available to study UTB SOI and DG devices. Drift Diffusion, Monte Carlo and Non-Equilibrium Green's Functions are considered and compared, with their relative strengths and weaknesses in terms of examining the impact of the body thickness fluctuations discussed in detail. This section provides the justification for the selection of a 3D Monte Carlo technique to conduct this research.

Chapter 4 describes the development of the Monte Carlo simulator used. Methods for the inclusion of quantum corrections, self-consistency and interface roughness are detailed, as well as other factors affecting the stability of simulations of these devices, such as the implementation of contacts, choice of mesh spacing and time step. The Drift Diffusion simulator used for initialization and comparison with the Monte Carlo simulator is also introduced and described.

In Chapter 5, the results from the various simulation studies using different versions of the Monte Carlo simulator are presented and discussed. The transport degradation and current variability from device to device are examined through statistical Monte Carlo studies and careful comparison to Drift Diffusion results.

Finally, in Chapter 6 the findings of this thesis are restated and summarised and suggestions are made for possible future work in terms of development of the Monte Carlo module.

Chapter 2

Scaling and Transport Phenomena of UTB Devices

2.1 Introduction

In comparison to conventional bulk MOSFETs, ultra thin body (UTB) silicon on insulator (SOI) and double gate (DG) MOSFETs offer superior electrostatics, but more complex transport behaviour.

At relatively large silicon body thicknesses (t_{Si}), these devices offer similar behaviour to their conventional counterparts. When scaled to intermediate dimensions (5 nm $< t_{Si} <$ 20 nm), a variety of effects such as volume inversion, increased phonon scattering, interface roughness and band splitting compete to influence transport. Below 5 nm, a rapid degradation of transport properties is experienced due to the impact of confined acoustic phonon and body thickness variation induced scattering.

This chapter first discusses the superior scaling properties of UTB SOI and DG devices in comparison with conventional bulk MOSFETs, before examining the

various mechanisms mentioned above that influence the transport of carriers in these architectures.

2.2 Scaling Properties of Conventional, UTB SOI and DG MOSFETs

Scaling of the conventional MOSFET into the nanometer regime (channel lengths sub 0.1 μm) has been an important area of research for many years now [29, 30, 31, 32], with a channel length of 25 nm being considered an achievable limit [33] through the employment of alternative materials and dielectrics. While devices with 15 nm gate length have been fabricated and demonstrated [34], the off and on state currents failed to meet desired specifications [27]. As the channel length shrinks, the impact of short channel effects (SCE) becomes increasingly pronounced especially at high drain voltage. The threshold voltage becomes dependent on both the channel length and drain voltage, and a phenomenon called 'punch through' is encountered at high drain voltage where the gate completely loses control of the channel, resulting in uncontrollable drain leakage current. All these effects are a manifestation of the drain induced barrier lowering (DIBL) which results in a direct reduction of the potential barrier between the source and drain by the drain potential and unwanted parasitic capacitance coupling between the drain and the channel. This in turn leads to a high subthreshold leakage current.

The usual approach for combatting these effects in conventional MOSFET structures is by increasing the doping within the channel, allowing for the depletion depth to be scaled along with the channel length, which reduces the electrostatic drain coupling and prevents the excessive growth of the subthreshold leakage current [35]. However, as the channel length is scaled down, the required doping increases considerably. The simple raising of the doping uniformly leads to a inappropriately high threshold voltage [36]. A more suit-

able option for short channel MOSFETs is to use retrograde doping, where a low-high (moving away from the oxide layer) doping profile is adopted, which allows for control of the depletion depth and reduction of SCE [36] without affecting adversely the threshold voltage. This method is used in epitaxial MOSFETs [30, 37]. Another option is to use high doping around the source and drain regions (referred to as 'halo' or 'pocket' doping) to prevent punch through, and thus allow for lower doping to be used in the middle of the channel for threshold voltage control. It has been shown that this method provides significantly better results in terms of controlling independently the threshold voltage and the short channel effects [33]. However, increasing the doping has adverse effects on transport and device performance due to increased impurity scattering and a higher transverse field which results in increased surface roughness scattering [38].

Additionally, the gate oxide thickness must scale with the channel length, to help maintain threshold voltage and keep control of the channel by the gate. However, the reduction of the oxide to thicknesses below 2 nm leads to an undesirably high gate tunneling leakage current [35].

Therefore, the limitations to scaling of conventional (bulk) MOSFETs come from tunneling through the gate oxide (gate leakage current) and the reduced device performance and band-to-band tunneling both associated with the increased doping used to control SCE.

By adopting a UTB SOI based architecture (illustrated in Fig. 2.1), the SCE can be controlled via the geometry of the device rather than by using high doping. Here, the thin body and the second oxide layer reduce the electrostatic coupling between the drain and the channel, lessening the amount of DIBL, thus helping to prevent SCE [39].

The two basic versions of single gate (SG) SOI MOSFET are the partially depleted (PD) SOI architecture where the silicon layer thickness is greater than the depletion layer, and fully depleted (FD) SOI architecture where the

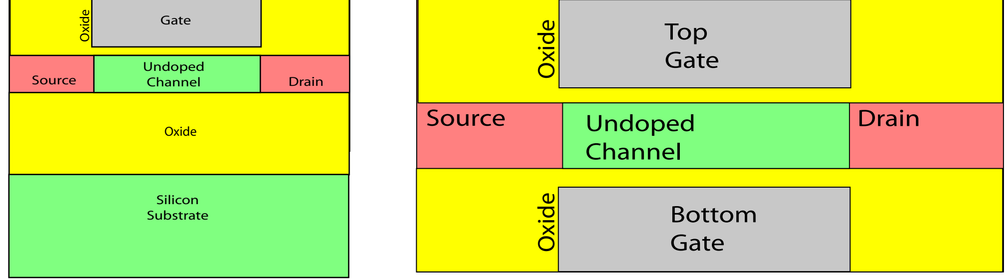

Figure 2.1: Schematics of SG-SOI (left) and DG (right) architectures, not shown to similar scales.

depth of the silicon layer is equal to the depth of the depletion region under the gate.

PD-SOI suffers a drawback in the floating body effect (FBE), where the region beneath the channel depletion layer can charge up from capacitive coupling to the gate and drain or from impact ionization in the drain region [39, 40]. This is undesirable as it can lead to unpredictable device behaviour. Connecting the floating body to the source to discharge it is one method of counteracting this problem, but this leads to slower switching times [35]. The scaling behaviour of the PD-SOI is similar to that of bulk MOSFETs and so it suffers from the same scaling limitation factors [1].

The FBE is not such a serious issue in FD-SOI as the entire silicon layer under the gate is fully depleted. Additionally, it tolerates much lighter doping compared to bulk or its PD counterpart [39]. FD-SOI therefore offers a more viable option in terms of scaling potential. The lightly doped channel results in less mobility degradation due to impurity scattering or high transverse fields, which can still be issues within PD-SOI where the threshold voltage and the SCE are controlled via channel doping. FD-SOI does present some drawbacks such as increased series resistance due to the thinness of the silicon layer [35]. Additionally, scaling of the oxide layer is still required for ever decreasing chan-

nel lengths, meaning gate leakage current is still an issue, although one that can be overcome via the use of high-κ dielectrics [35]. Also, this architecture does suffer from self heating as the back oxide layer does not efficiently conduct heat away [40]. A significant drawback of FD-SOI compared to PD-SOI, or even bulk MOSFETs is the difficulties in controlling the thickness of the uniformity of the silicon body in extremely scaled devices.

Another promising device structure is the DG MOSFET illustrated in Fig. 2.1, where in addition to the second oxide layer, there is a second gate, which itself creates an inversion layer and a channel through which carriers flow. It has been shown that this architecture improves scaling potential compared to the SG-SOI MOSFET described above [41, 42, 43]. The second gate helps improve SCE while maintaining a good subthreshold slope (no greater than 80 mV/decade is desirable), which becomes a problem for aggressively scaled SG-SOI devices [39]. It has been predicted that DG architectures could be almost twice as scalable as their bulk predecessors [39, 44]. Experimental examination of a UTB DG MOSFET has recently been carried out [45], showing a good subthreshold slope (76 mV/dec), as well as mobility enhancement, a significant improvement when compared to SG operation. Again, in DG transistors the transverse electric field is reduced further compared to conventional bulk and SG-SOI MOSFET designs.

In both SG-SOI and DG devices, as the geometry of the device controls the SCE, low (intrinsic) channel doping can be used, thus preventing degradation of transport through increased impurity scattering. Additionally, subthreshold and the band-to-band tunneling leakage current can be controlled via the thinness of the silicon layer [35].

A useful illustration of how alternative device architectures cope with SCE has been presented by Skotnicki [46, 47]. By quantifying DIBL using equation (2.1), conventional, SG-SOI and DG MOSFETs can be compared.

$$DIBL = 0.80\frac{\varepsilon_{Si}}{\varepsilon_{ox}}\left(1 + \frac{X_j^2}{L_{eff}^2}\right)\frac{t_{ox}}{L_{eff}}\frac{t_{dep}}{L_{eff}}V_D \tag{2.1}$$

Here, L_{eff} is the effective channel length, X_j is the depth of the source and drain regions, t_{ox} and t_{dep} are the thickness of the oxide and depletion layer respectively, ε_{ox} and ε_{Si} are the permittivity of the oxide and silicon respectively and V_D is the applied drain bias.

In all cases, $\varepsilon_{ox} = 3.9$, $\varepsilon_{Si} = 11.7$ and V_D assumed to be 1 V. Additionally, L_{eff} is taken to be $2/3L_g$, where L_g is the gate length. Finally, t_{dep} is assumed to be the same as X_j, and t_{ox} to be 1/30 of L_g. The equation also highlights how important it is to scale all parameters consistently, as scaling L_{eff} without scaling the other parameters would lead to an overall increase in DIBL. These scaling assumptions were given by Skotnicki as a good approximation for reasonable device design rules.

For a conventional MOSFET architecture, t_{dep} and X_j are assumed to equal half of L_g. Inserting this into (2.1) gives a value of 140 mV for DIBL in bulk MOSFETs. For a SG-SOI MOSFET, t_{dep} is now assumed to be equal to the silicon thickness (i.e. fully depleted), which itself is assumed to be 1/3 of L_g. This gives a value of 75 mV for DIBL. Finally, a DG MOSFET where t_{dep} is half of the silicon thickness (treating each gate's depletion region separately), with the silicon thickness being the same as that for the SG-SOI example will have 32 mV DIBL which is less than half of that calculated for SOI and a quarter of that for a conventional MOSFET, highlighting the DG MOSFETs superiority in efficiently controlling SCE purely by geometry. The above calculations are summarised in Table 2.1.

	Conventional	SOI	DG
L_{eff}	2/3	2/3	2/3
X_j	1/2	1/3	1/6
t_{ox}	1/30	1/30	1/30
t_{dep}	1/2	1/3	1/6
DIBL (mV)	140	75	32

Table 2.1: Comparison of DIBL in various MOSFET structures based on (2.1). All dimensions are expressed as fractions of the gate length, L_g.

While further variations of multiple gate architectures (triple or even quad gates) may offer even better performance [41], the difficulties encountered in fabrication of these devices are a significant drawback. Acceptable performance can be obtained via FD-SOI and DG MOSFET designs, and they remain viable options in terms of fabrication. The International Technology Roadmap for Semiconductors (ITRS) [27] suggests SOI taking over from bulk architectures around the 45 nm technology node, with DG coming to prominence soon after, both near term targets.

The rest of this chapter comprises an examination of various quantum mechanical and transport aspects of the operation of UTB SOI and DG MOSFET devices, focusing on phenomena that becomes important when the silicon thickness is scaled well below 20 nm.

2.3 Volume Inversion

Arguably one of the most prominent effects associated with the operation of DG UTB MOSFETs is that of volume inversion, first observed by Balestra *et al.* [48].

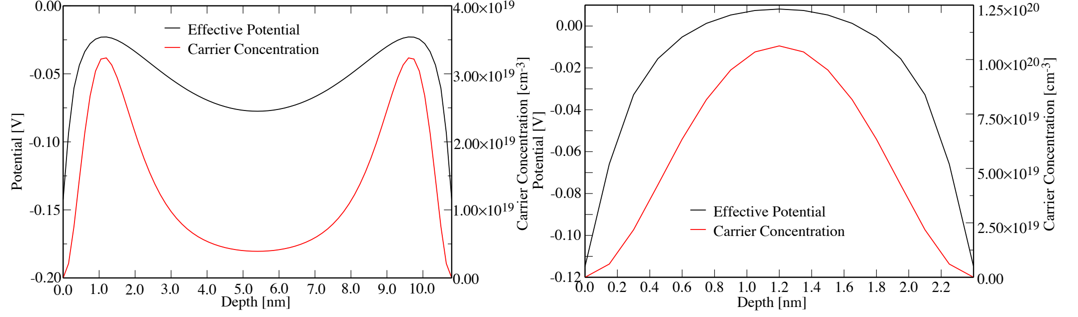

Figure 2.2: Comparison of potentials and electron distributions in DG MOSFETs with t_{Si} = 10.8 nm (left) and t_{Si} = 2.4 nm (right). These plots were generated during this work, using the Drift Diffusion simulator with Density Gradient quantum corrections described in a later chapter.

In DG MOSFETs, both gates can create an inversion layer near each of the semiconductor/oxide interfaces with the appropriate bias applied. In effect, this is like two MOSFETs with a single drain, substrate and source [49].

In the case when the thickness of the silicon layer is greater than the sum of the two depletion regions, the device operates as two MOSFETs in parallel, with two channels, one towards each interface (see Fig. 2.2).

For sufficiently thin silicon layers, with suitable gate voltages applied, due to quantum confinement the entire volume of silicon can be forced into strong inversion (with the depletion regions meeting, and carriers no longer confined towards the interface, instead spread through the whole body and peaking in the centre), leading to the 'volume inversion' effect [49, 50]. This is also illustrated in Fig. 2.2, which also highlights that this is a quantum mechanical effect, with the peak of the electron concentration moved away from the semiconductor/insulator interface, in contrast with the classical carrier distribution which peaks at the interface [51].

The benefits associated with this phenomena include [48, 49] an increase in the number of carriers in the channel and a decrease in the impact of rough-

ness and fixed charges at the interfaces. These two effects ultimately produce an increase in both the current and transconductance. Additionally, volume inversion reduces hot carrier effects, and produces a steep threshold slope [52]. The transverse electric field is lower under volume inversion conditions, which reduces the interface roughness scattering and increases the carrier mobility.

The conditions defining volume inversion have an impact upon the other significant transport phenomena (scattering associated with phonons, Coulomb interactions and interface roughness) observed in UTB SOI and DG MOSFETs. A discussion of these mechanisms and their resulting effect follows in this chapter.

2.4 Enhanced Phonon Scattering

In ultra thin silicon layers, such as the one employed in UTB SG-SOI and DG MOSFETs, an increase in phonon scattering is observed which contributes to transport degradation. Two phenomena contribute to this: confinement of electrons in thin silicon layers leading to more bulk phonons available to participate in transitions between electron states [53]; and the confinement of acoustic phonons in thin silicon layers [54].

The first phenomena that increases the phonon scattering results from greater confinement of carriers in thin t_{Si}, which leads to a reduction of uncertainty on the electrons' positions perpendicular to the interface, allowing for a wider distribution of values for the electrons' momenta. This provides more phonons to contribute to transitions between electron states, leading to an increase in phonon scattering [53].

Shoji *et al.* [55, 56] considered the phonon limited mobility in UTB SOI and DG MOSFETs. The effect of band splitting, when carriers populate different

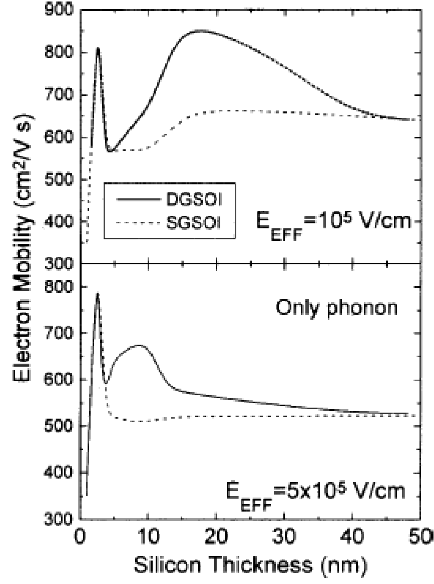

Figure 2.3: Phonon limited mobility versus t_{Si} from [49], comparing trends in DG and SG-SOI MOSFETs.

subbands based on their energy, becomes important. This effect is discussed in more detail in the next section.

Phonon scattering starts to produce an appreciable decrease in mobility for higher energy subbands for $t_{Si} < 20$ nm, and for the lowest energy subbands for $t_{Si} < 10$ nm. Therefore, a total drop in phonon limited mobility is not observed until $t_{Si} < 10$ nm due to the higher occupation of the lowest energy subbands.

The combination of these factors allow Shoji [56] and Gàmiz [49] to observe three t_{Si} defined phonon limited mobility regimes, as can be observed in Fig. 2.3. For large values of t_{Si}, the mobility of DG is similar to that observed in SG SOI MOSFETs, as the inversion layers are kept separate. As t_{Si} decreases, and the DG MOSFET enters the volume inversion regime the mobility trends of the two begin to differ as the inversion layers interact. A peak

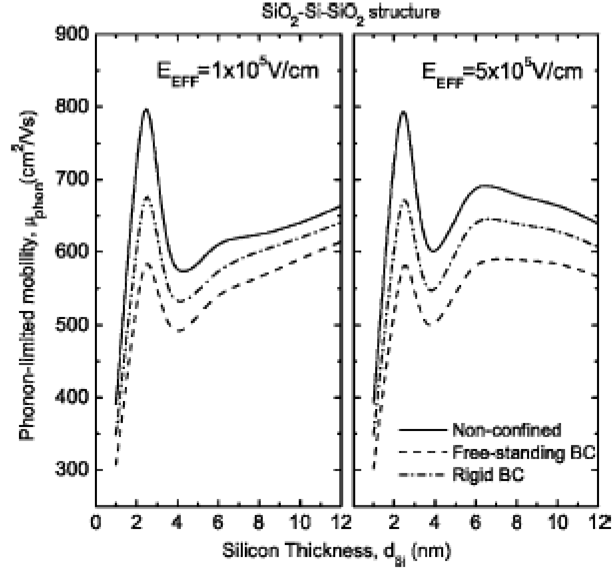

Figure 2.4: Dependence of phonon limited electron mobility on silicon layer thickness from [57], showing the significant degradation coming from confined acoustic phonon scattering. The three different curves, refer to boundary conditions used by the authors when including the phonon scattering mechanism. They suggest that the most appropriate boundary condition will vary depending on the material and structure under consideration.

value of mobility comes around $t_{Si} = 10$ nm for the reasons mentioned above. Below $t_{Si} = 10$ nm, the geometrical confinement leads to an overall increase in phonon scattering in both structures. The increase in mobility observed in this region is due to band splitting and will be discussed in the next section.

The second aspect of the increased phonon scattering is associated with the confinement of acoustic phonons in thin silicon layers which results in a modification of phonon modes where there is a mismatch between the dielectric constants of the silicon and the oxide layers [57]. This phenomena has been

observed experimentally [58], and produces a significant degradation in electron mobility.

The increase in confined phonon scattering results in a further degradation of electron mobility down to $t_{Si} = 2$ nm, as shown in Fig. 2.4. The most significant impact is in the region 5 nm $< t_{Si} <$ 10 nm, where the transport enhancement associated with volume inversion is counteracted [59].

The importance of phonon scattering comes from the fact that the mechanism is coupled to the device itself, so no matter how perfect the fabrication process, the limitation will always be there at room temperatures (though is avoidable at low temperatures) [53]. The introduction of high-κ dielectric will only exacerbate these effects by introducing additional soft optical phonon scattering [21, 60].

2.5 Band Splitting

Scaling of devices to ever smaller dimensions ultimately leads to an increased impact of quantum effects upon transport and device characteristics [44, 51, 62, 63, 64]. In thicker silicon layers, confinement is a product of an external bias, however, at thinner silicon layers the confinement is a geometrical effect resulting from the close proximity of the layer boundaries. This size induced quantization shifts the peak concentration of carriers away from the interface (contrasting with the classical distribution), additionally increasing the effective oxide thickness [50, 65].

As reported in [66], and shown in Fig. 2.5, DG MOSFETs experience a mobility enhancement when scaled to body thicknesses between approximately 3 and 5 nm. This enhancement is attributable to subband splitting and the carrier occupation of 2-fold (or unprimed subband) valleys that have a lower effective mass compared to the carriers in the 4-fold (or primed subband) valleys in the

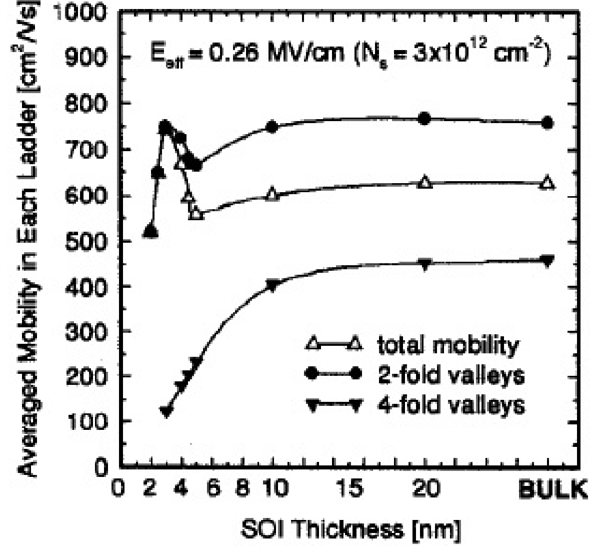

Figure 2.5: Mobility in different subbands and overall mobility, from [61]. Shows the increase in mobility in the 3-4 nm region in the 2-fold valleys, that contributes to the increase in the overall mobility.

direction of current flow, as described in Table 2.2 and shown in Fig. 2.6. This subband structure engineering can be exploited in strained silicon MOSFETs, as well as the UTB SOI MOSFETs (SG or DG) considered here [61].

In UTB SG-SOI and DG MOSFETs, the dependence of mobility on silicon thickness can be explained as a result of the energy difference between the 2-fold and 4-fold valleys. When the t_{Si} is thick (greater than 10 nm), the top and bottom interfaces are separated so that the band structure is the same as in bulk MOSFETs and the mobility is constant.

The differences in the confinement direction effective mass (higher in the 2-fold valleys) and the inversion layer (thinner in the 2-fold valleys) leads to a split in occupation as the silicon layer reduces. The thickness of the inversion layer for the 4-fold valleys means that the size of the silicon layer influences their ground state energy (E_0) first when $t_{Si} < 5$ nm, pushing it upwards and creating a

Unprimed (2-fold)	Primed (4-fold)
lower conductivity mass	higher conductivity mass
($0.19m_0$)	($0.315m_0$)
higher mass normal to transport	lower mass normal to transport
($0.916m_z$)	($0.19m_z$)
higher mobility	lower mobility
lower subband energy	higher subband energy

Table 2.2: Description of the subband structure for (100) Si orientation.

difference in E_0 between the 2- and 4-fold valleys (ΔE_0) and is depicted in Fig. 2.7. Carriers begin to occupy the 2-fold valleys in greater numbers and this leads to an overall increase in mobility due to the lower conductivity mass [52, 61]. This trend continues till $t_{Si} < 3$ nm and all the electrons occupy the 2-fold valleys. Now the size of the silicon layer is of a similar, or smaller, size to the inversion layer, leading to an increase in scattering from acoustic phonons and a resulting decrease in overall mobility [67]. Fig. 2.5 shows the mobility dependence on silicon thickness in the 2-fold and 4-fold valleys, as well as the total mobility. This effect can only be captured in quantum mechanical simulations as classically all valleys are equally populated [51].

Taking into account the impact of the transverse effective field, the relationship between t_{Si} and mobility is less straightforward. In the region 20 nm $> t_{Si} >$ 5 nm, mobility increases with a decreasing electric field. However below $t_{Si} = 5$ nm, the dependence of mobility on field lessens, and vanishes when $t_{Si} < 3$ nm with mobility independent of t_{Si} [61, 67] .

Further, the impact of temperature has been investigated, showing that at low temperatures there is a greater improvement in mobility for UTB DG MOSFETs compared to that observed at room temperature [68].

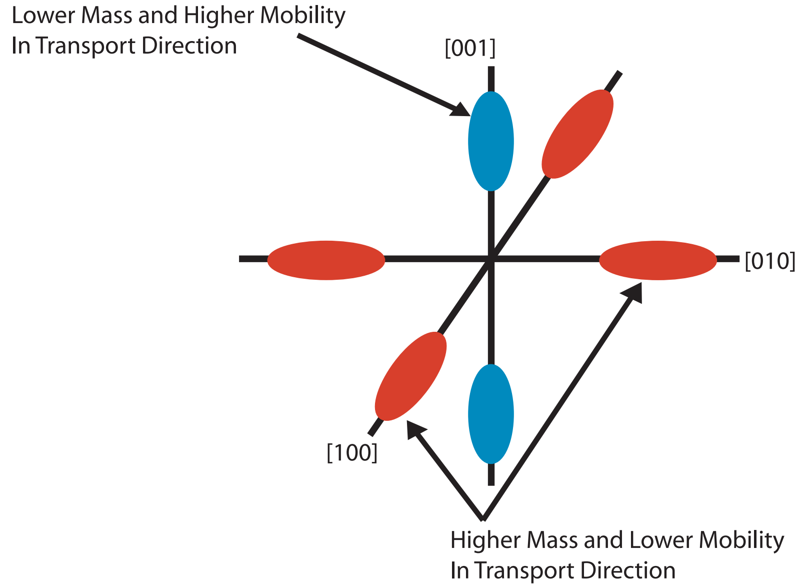

Figure 2.6: Subband structure, showing 4- (in plane) and 2- (out plane) fold valleys in red and blue respectively.

2.6 Coulomb Scattering

Typically Coulomb scattering depends on several factors, such as the distribution of electrons in the inversion layer, the distribution of charged scattering centres, screening of the charged centres by mobile carriers, the charged centres' correlation and image charges. In DG MOSFETs, the carrier distribution differs from that in bulk MOSFETs hence the screening, and the relative position between the carriers and the charged centres will also differ, impacting upon the effect of the Coulomb potential screening on electron mobility [69].

Coulomb scattering in UTB SOI MOSFETs results from carrier interactions with trapped charges at the two Si/SiO_2 interfaces, as channel doping in both SG SOI and DG is low enough that Coulomb scattering due to dopant impurities can be considered negligible [70]. It becomes the dominant mechanism at lower inversion charge densities (compared to phonon scattering discussed pre-

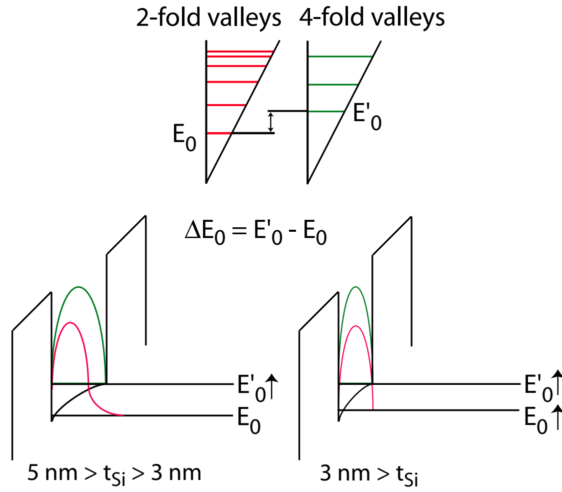

Figure 2.7: Comparison of 2- and 4-fold valleys showing the energy separation that leads to splitting, and highlighting the confinement from both the field and structure and its influence on the ground state as t_{Si} decreases. Adapted from [67]

viously), though an increase in the effective vertical field and the corresponding increase of the carrier concentration in the channel reduces the impact of the charged centres due to screening [49].

Uchida *et al.* [71] experimentally observed the impact of Coulomb scattering, showing for SG SOI MOSFETs a Coulomb limited mobility degradation is evident for thinner silicon layers when compared to thicker ones. Additionally, he observed a suppression of Coulomb scattering in DG MOSFETs compared to SG SOI MOSFETs due to an increase in screening. This is associated with the volume inversion effect. At thinner silicon layers, volume inversion improves the screening of the charged centers, thus reducing the impact of Coulomb scattering.

Coulomb scattering is, however, considered to be an issue which could be overcome via technological advancement in the device's fabrication [53].

2.7 Interface Roughness and Body Thickness Variation

Roughness patterns at the top and bottom Si/SiO_2 interfaces induce two scattering mechanisms that influence transport in UTB SOI MOSFETs: interface (or surface) roughness scattering and body thickness fluctuations.

Numerous authors have considered the impact of surface roughness scattering in DG MOSFETs, and found the resulting transport degradation to be as significant as in bulk MOSFETs, if not more so.

In the study by Kathawala *et al.* in [50], Ando's model for surface roughness scattering [51] is used in a quantum corrected Monte Carlo simulator. For a device with t_{Si} = 10 nm, the comparison of simulations with and without interface roughness scattering shows a reduction in current and mobility in the presence of this additional scattering mechanism. Bufler *et al.* [72] carried out classical self-consistent Monte Carlo simulations using a specular/diffusive surface roughness scattering model [73] to examine this mechanism in DG MOSFETs with t_{Si} = 6.25 nm and found the additional scattering to have an effect upon the on-current improvement expected in these structures.

Gàmiz [74] examines the role of surface roughness scattering more rigorously. The mobility for a range of transverse effective fields for different t_{Si} was simulated. At low fields, degradation is only observed when t_{Si} < 5 nm. This is due to the fact that the confinement at low fields is determined by the thinness of the silicon layer. At high fields, degradation is observed for all thicknesses as the field is forcing carriers toward the interfaces [75].

Comparisons of the impact of surface roughness scattering on UTB MOSFETs with both the single and double gate configuration have also been carried out by the same authors [74]. Three regions of surface roughness limited mobility are observed. The first, at thick silicon layers (t_{Si} > 20 nm), where there is

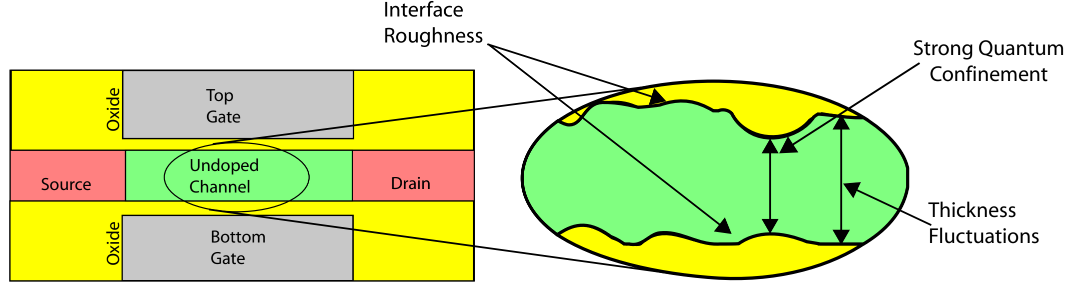

Figure 2.8: Source of body thickness fluctuations resulting from roughness patterns at the top and bottom interfaces.

no interaction between inversion layers in the DG device, the mobility curves coincide with the universal mobility value. For 3 nm $< t_{Si} <$ 20 nm, the DG device enters the volume inversion regime, which leads to an increase of mobility in comparison to the SG SOI MOSFET. Below $t_{Si} = 3$ nm, the mobility in the two architectures again coincide due to other scattering mechanisms (most notably phonon scattering) dominating, resulting in an overall drop in mobility. Additionally, the presence of the second oxide layer in both configurations produce similar geometrical confinement effects.

The studies of Gàmiz [49] used quantum corrected Monte Carlo simulations [76] with a modified version of the roughness scattering model of Ando [51] to demonstrate the degradation in transport for this mechanism. His explanation was that at smaller t_{Si} the presence of a second interface, and the accompanying additional surface roughness scattering, was responsible for the drop in performance.

However, the drop in mobility for $t_{Si} <$ 5 nm can be explained more accurately by examining the impact of body thickness fluctuations caused by the unique roughness patterns at the top and bottom of the silicon layer (see Fig. 2.8).

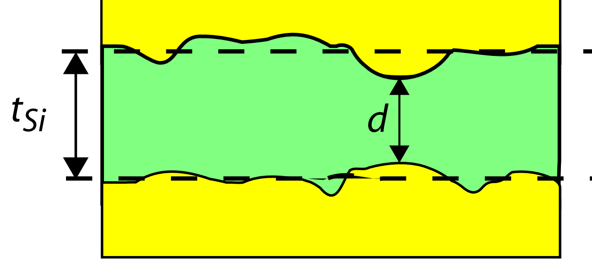

Figure 2.9: The relationship between the nominal thickness of the silicon layer (t_{Si}) and the roughness limited actual silicon thickness ($d = t_{Si} - \Delta(r)$ where $\Delta(r)$ accounts for the contribution of both oxide layers).

These thickness fluctuations lead to variation in subband energies [71, 77] which results in a shifting of the ground state which may be described by [28]:

$$E_0 = \frac{h^2}{8m^* t_{Si}^2} \tag{2.2}$$

Where h is Plank's constant and m^* is the effective mass of the electron. The height of the interface roughness fluctuations (Δ), is usually assumed to be $\pm$ 1 atomic layer, so with 2 interfaces, the total fluctuation is 4 atomic layers [28]. Hence, as t_{Si} varies along the channel, so does E_0. The shifting of the ground state leads to the variation in subband energies as mentioned above. The resulting fluctuations in potential is given by [28, 77]:

$$\Delta V = \frac{\partial E_0}{\partial t_{Si}} \Delta = -\frac{h^2}{4m^* t_{Si}^3} \Delta \tag{2.3}$$

From (2.3), the fluctuations are dependent on the height of the roughness, so for smoother surfaces, there is less fluctuation, and hence less scattering.

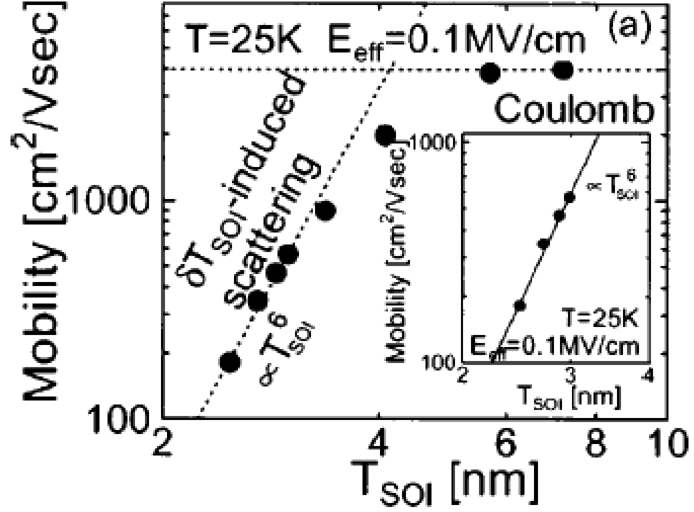

Figure 2.10: Mobility dependence on t_{Si} from [28], showing the relationship stated in (2.4). Study carried out at low temperature to suppress effects from increased phonon scattering at these dimensions.

When $t_{Si} < 4$ nm, the fluctuation in conduction-band energy from the shift in ground state exceeds thermal energy at room temperature leading to the potential barriers that act as an additional scattering potential for electrons in the channel [66].

The relationship between mobility and silicon thickness has been shown to be [28, 78, 79]:

$$\mu \propto \left(\frac{\partial E_0}{\partial t_{Si}} \Delta \right)^{-2} \propto t_{Si}^6 \tag{2.4}$$

and is demonstrated in Fig. 2.10. This explains how the reduction of t_{Si} leads to a dramatic degradation of μ.

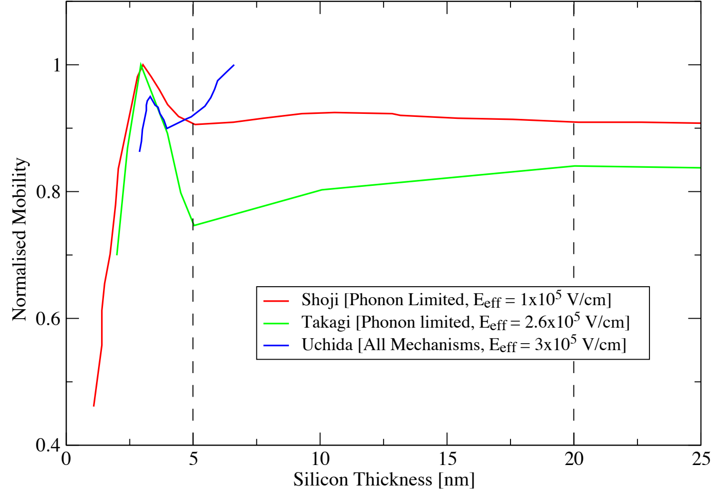

Figure 2.11: Comparison of mobility vs. silicon thickness trends in SG-SOI MOSFETs from Shoji [55], Takagi [67] and Uchida [66]. The differences above $t_{Si} = 5$ nm are attributable to differences in field and the scattering mechanisms employed in the different studies.

Scattering from body thickness fluctuations becomes prominent when $t_{Si} <$ 4 nm, and, along with confined acoustic phonon scattering, leads to a sharp drop in the mobility observed in Fig. 2.5 [66]. Hence, this scattering mechanism imposes a severe limitation on carrier mobility as the silicon thickness in UTB SOI and DG MOSFETs is decreased.

The uniqueness of the roughness pattern from interface to interface leads to variation in drive current from device to device that may hamper devices' integration in future circuits.

2.8 Summary

Considering all the mechanisms discussed in this chapter, three distinct areas of operation for UTB MOSFETs can be identified, each defined by silicon

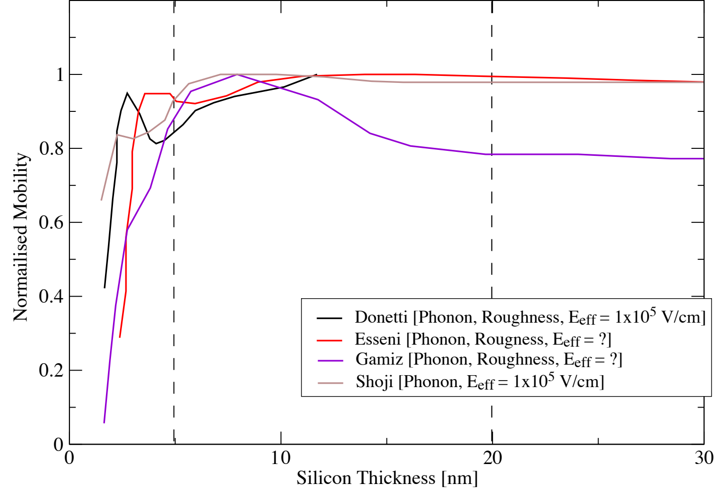

Figure 2.12: Comparison of mobility vs. silicon thickness trends in DG MOSFETs from Donetti [57], Esseni [77], Gàmiz [74] and Shoji [56]. Differences in location of peak mobility is attributable to variations in scattering phenomena captured.

layer thickness. These three regions are identified in the work of Gàmiz [49], Uchida [71] and Venkatesan [80] and can be recognised in Fig. 2.5 and summarised in Fig. 2.11 for SG-SOI and Fig. 2.12 for DG MOSFETs.

For $t_{Si} > 20$ nm there is little or no improvement in double gate MOSFETs in comparison to their SG SOI or bulk counterparts. This is a result of electrons behaving in a similar manner in all three architectures with the subband structures coinciding. This has been demonstrated by numerous authors [49, 55, 56, 70, 77].

For 20 nm $> t_{Si} >$ 5 nm, the reduction in t_{Si} leads to UTB DG MOSFETs entering the volume inversion regime, and an appreciable improvement in mobility can be observed. The interaction of the top and bottom inversion layers alters the subband structure, and in conjunction with the spread of electrons through the silicon layer, the impact of several scattering mechanisms is re-

duced. In particular channel orientations a peak in mobility can be observed due to band splitting.

Finally, for $t_{Si} < 5$ nm, the presence of the second SiO_2 layer for both DG and SG SOI MOSFETs leads to the mobility curves for both coinciding and dropping sharply as confined acoustic phonon and body thickness variation scattering starts to dominate.

Ultimately, these effects culminate to impose dimensional limitations in the scaling of UTB SOI MOSFETs with either one or two gates.

While some of these problems may be eventually overcome via improved fabrication (such as the Coulomb scattering effect and scattering from body thickness fluctuations), others are inherent to the device structure and materials used (such as scattering from phonons) and will degrade transport regardless of fabrication quality.

The focus in this study is the device performance variations from differences in body thickness dominated scattering due to the unique interface and body thickness pattern in each nano UTB SG-SOI or DG MOSFET.

Chapter 3

Simulation Methodology

3.1 Introduction

In order to capture the device variability phenomena discussed in the previous chapter, a simulation approach is required that can take into account quantum mechanical effects. In this chapter, the strengths and weaknesses of Drift-Diffusion (DD), Monte Carlo (MC) and Quantum Transport based on Non-equilibrium Green's Function (NEGF), three generic simulation approaches capable of including quantum effects, are discussed and compared.

The Boltzmann transport equation (BTE) (3.1) describes semiclassical carrier transport.

$$
\begin{aligned}
&\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_r f + \frac{e\mathbf{F}}{\hbar} \cdot \nabla_k f \\
&= \sum_{k',\lambda} S_\lambda(\mathbf{k},\mathbf{k}') f(\mathbf{r},\mathbf{k}',t)(1 - f(\mathbf{r},\mathbf{k},t)) \\
&- S_\lambda(\mathbf{k},\mathbf{k}') f(\mathbf{r},\mathbf{k},t)(1 - f(\mathbf{r},\mathbf{k}',t))
\end{aligned}
\tag{3.1}
$$

It determines the distribution function f which is formulated in a seven dimensional phase space, consisting of position $\mathbf{r}$ and momentum $\mathbf{k}$ at time t, along with velocity $\mathbf{v}$ and force $\mathbf{F}$. $S_\lambda(\mathbf{k}, \mathbf{k}')$ is the scattering probability from state $\mathbf{k}$ to $\mathbf{k}'$. The right hand side is the collision integral (often represented as $(\partial f/\partial t)_{coll}$), describing the scattering effects. The direct solution of the BTE is cumbersome [81], so alternate approaches have been derived, including DD and MC.

In order to couple properly the current flow to the field distribution of the simulated devices, it necessary to use self-consistent simulations [82]. This is implemented in all 3 simulation techniques by solving Poisson's equation (3.2) in order to calculate the potential resulting from the mobile charge distribution present:

$$\nabla \cdot (\varepsilon \nabla \psi) = q(n - p + N_A - N_D) \tag{3.2}$$

Here, ε is the permittivity of the material, ψ the potential, q the charge, n and p are the mobile charges, and N_A and N_D the acceptor and donor concentrations. The three simulation approaches use different methods to update the mobile charge distribution based on the updated potential. These two operations are carried out self-consistently until a specified convergence is reached.

In DD, Poisson's equation is solved self consistently with the current continuity equation. In MC, a statistical ensemble of carriers is propagated within the device suffering random scattering mechanisms. In quantum transport, the Schrödinger equation is solved with open boundary conditions using the Green's Function formalism.

The necessity to capture quantum effects increases as the dimensions of the simulated device shrink [82]. Many of the phenomena affecting transport in UTB SOI MOSFETs, as described in the previous chapter, require a quantum mechanical approach in order to be modelled accurately.

While NEGF inherently describes the quantum effects, with DD and MC it is necessary for quantum corrections to be implemented. A range of various methods have been employed including Schrödinger [83], Wigner [84], Bohm [85] and Effective Potential [86] based techniques.

Typically, these corrections are added to the classical potential calculated via the solution to Poisson's equation, with the resulting quantum potential used to update the current continuity equation in DD, or to calculate the driving force for the particles in MC. Validation and/or calibration is then carried out via careful comparison to solutions using a NEGF algorithm or Poisson-Schrödinger solvers [65]. The use of a quantum correction allows for the computational efficiency of these methods to be retained while making the capture of quantum effects, such as size quantization and, with lesser success, tunneling, possible.

In order to effectively study MOSFET variability, it is necessary to capture the responsible mechanisms including quantum confinement variation and the corresponding scattering which varies from device to device. Hence, in order to be useful, all the methodologies discussed in this chapter must be capable of capturing the electrostatic and/or transport limiting scattering effects that lead to variability in devices.

3.2 Drift-Diffusion

The DD formalism [87, 88, 89] in unipolar devices like MOSFETs, involves self consistently solving the Poisson (3.2) and Current Continuity (3.3) and (3.4) equations (for electrons in n-channel devices), until convergence in the current solution is met.

$$\nabla \cdot J_n = qR \tag{3.3}$$

$$J_n = J_{diffusion} + J_{drift} = qD_n\nabla n - q\mu_n n\nabla\psi \tag{3.4}$$

Where J_n is the electron current density, R the net recombination rate, $J_{diffusion}$ and J_{drift} the current densities for diffusion and drift respectively and D_n and μ_n the diffusion coefficient and mobility for electrons. The other symbols have the same meaning as before.

The equation for current density (3.4), which combines drift and diffusion components, can be derived from the first moments of the BTE (3.1) [87]. It represents an approximation of the BTE, with several assumptions and simplifications applied: the use of the relaxation time approximation, a slowly varying field at any point in the device, and equal lattice and carrier temperatures [87].

DD provides a relatively simple, computationally efficient simulation basis that is easily extendable into two- and three-dimensional domains [90]. It has been at the heart of device simulation for many years, and provided an extremely good tool for modeling devices up to entry into the nanometer regime. The necessary conditions are to assume that the problem is localised, with carriers having no 'memory' of the numerous successive collisions they undergo while traversing the device and that the scattering is an additional 'frictional' force acting upon the current flow [91].

When simulating modern decananometer scale MOSFETs it is most applicable in the subthreshold regime, where the current is not strongly coupled to the solution of Poisson's equation and the electrostatics dominate. In this instance the source to drain potential barrier is sufficiently high so that little mobile charge inhabits the channel. The barrier is controlled by both V_G and V_D and by sources of barrier fluctuation such as line-edge roughness (LER), interface roughness and variations in dopant placement [37, 92, 93]. Above threshold, it has been shown that DD fails to properly describe the on-current (I_{on}) and its variability [94, 95, 96, 97].

Scaling of devices leads to a breakdown in the assumptions used to derive the DD equations. The problem becomes non-equilibrium and 'non-local' and the previous scattering history has to be taken into account.

The classical version of DD presented above does not deal with non-equilibrium transport, as carriers are treated as being in thermal equilibrium with the lattice, so carrier heating is not taken into consideration [36, 98]. As devices are scaled the electric field increases resulting in more hot carriers, so the validity of DD simulations decreases - carriers can now gain kinetic energy greater than the thermal energy.

As DD responds immediately and locally to changes in the electric field, it cannot capture non-local effects. Rapid changes in field occur in small devices, where μ_n and D_n are no longer simply connected to the local field, and become dependent on the carrier history [91].

In small devices, carriers can exceed the saturation velocity, resulting in velocity overshoot [99, 100], a non-equlibrium phenomena which is not accounted for in DD [87]. As carriers pass from a low to high field region, the average velocity reacts quicker to the change than the average kinetic energy, so the instantaneous mobility remains at its previous higher level. This results in a spatial velocity overshoot. As the carriers propagate further into the channel, the velocity returns towards the value that might be expected in a long channel device. DD assumes a simple field dependent mobility model, locally coupled to the electric field, so this effect is not captured [94, 101, 102]. It has been demonstrated that velocity overshoot is a desirable quality, increasing transconductance [103] and improving the switching time in circuits [104] making it a relevant area of research.

Methods to extend the applicability of the DD formalism have been explored, using more advanced mobility models [105] that do extend the validity of the model. Extension to higher moments of the BTE (such as the hydrodynamic or energy transport models [106]) where the thermal equilibrium assumption

is no longer applied [72] are also in common use, yet discrepancies still appear when compared to more sophisticated transport models.

Therefore, as devices scale to decananometer dimensions, methods based on simplifications of the BTE become less appropriate, and it is likely that more direct solutions of the BTE will become necessary [107].

3.2.1 Quantum Corrections in Drift-Diffusion

To further extend the validity and usefulness of the DD simulation approach, the introduction of a quantum corrected potential has allowed for effects relating to confinement and tunneling to be captured [82].

The Density Gradient formalism [108] is derived, similarly to the Bohm interpretation of quantum mechanics [109], as a correction to the classical potential, based on a modification to the BTE using the Wigner distribution function (3.5).

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_r f + \frac{1}{\hbar}\mathbf{F}^{\mathbf{QC}} \cdot \nabla_k f = \left(\frac{\partial f}{\partial t}\right)_{coll} \tag{3.5}$$

Here $\mathbf{F}^{\mathbf{QC}}$ is the quantum correction force defined as a derivative of the sum of the classical and quantum correction potentials (3.6).

$$\mathbf{F}^{\mathbf{QC}} = -\nabla\left(\psi_c - \frac{\hbar^2}{12m^*}\nabla^2 \ln(n)\right) \tag{3.6}$$

Here, ψ_c is the classical potential, m^* the effective mass and n the electron concentration. Writing the equilibrium carrier concentration in terms of the correction to the potential, then expanding and taking the lowest order com-

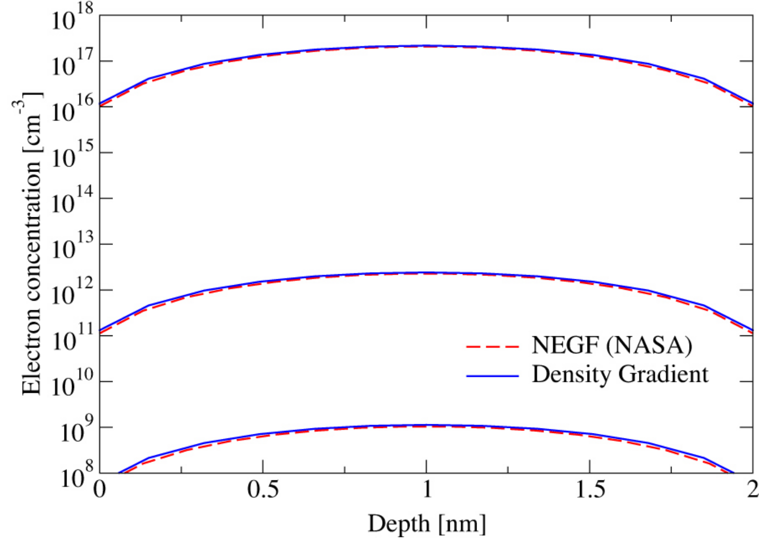

Figure 3.1: Comparison between vertical electron distribution (i.e. the quantization direction) for NEGF and Density Gradient results for various V_g in a DG MOSFET with $t_{Si} = 2$ nm, showing a good agreement between the two methods [110].

ponent gives an extra quantum correction dependent on the second derivative of the carrier concentration which can be added to the current equation giving:

$$J_n = qD_n \nabla n + q\mu_n n \nabla \psi + 2\mu_n \nabla \left(b_n \frac{\nabla^2 \sqrt{n}}{\sqrt{n}} \right) \tag{3.7}$$

Where $b_n = \hbar^2/(4qm^* r)$ and r is a fitting parameter, defined between 1 and 3 depending on the number of subbands filled [111]. Expressing the current in terms of the quasi-Fermi level gradient leads to an equation for the quasi-Fermi level [112]:

$$\phi_n - \psi_c + \frac{k_B T}{q} \ln \frac{n}{n_i} = 2b_n \left(\frac{\nabla^2 \sqrt{n}}{\sqrt{n}} \right) \tag{3.8}$$

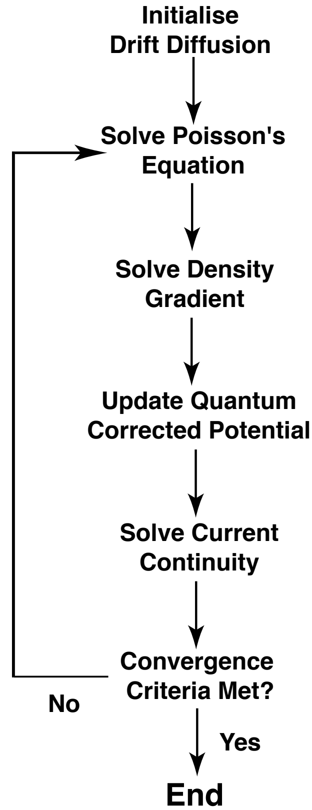

Figure 3.2: Simulation flow for Drift Diffusion using a modified Gummel algorithm

Where ϕ_n is the quasi-Fermi level. The parameter b_n can be calibrated to reproduce the vertical electron distribution obtained from a fully quantum NEGF simulation, as can be seen from Fig. 3.1. It has been shown that the Density Gradient corrections can capture accurately carrier confinement and mimic source-to-drain tunneling via lowering of the potential barrier in 3D simulations [110, 113, 114]. The dependence on the gradient of the carrier density also lends the method an element of non-locality. A typical simulation flowchart for a DD simulator employing Density Gradient quantum corrections is shown in Fig. 3.2.

3D DD simulations employing Density Gradient corrections have been used to study body thickness fluctuations in ultra thin body (UTB) single gate silicon on insulator (SG-SOI) MOSFETs [115]. This method successfully captured the electrostatic and confinement effects, and the accompanying shift in V_T and reduction in I_{on} [36], but it failed to address the additional quantum confinement scattering variation and the associated increased I_{on} variability.

An alternative quantum correction method used in DD is the Effective Quantum Potential approach [116], which takes the form:

$$\psi_{effective}(x) = \int \psi_c(x_0) G((x_0 - x), a_0) dx_0 \tag{3.9}$$

$$G((x_0 - x), a_0) = \frac{1}{a_0\sqrt{2\pi}} \exp\left(\frac{-(x_0 - x)^2}{2a_0^2}\right) \tag{3.10}$$

$$a_0^2 = \frac{\hbar^2}{12 m^* k_B T} \tag{3.11}$$

Where (3.10) is a Gaussian distribution used to smooth the classical potential in (3.9). Here, a_0 is the standard deviation and is effectively the quantum mechanical size of the particle [86], and can be tuned by varying m^* [117, 118]. This is based on an approach developed by Feynman [119] taking the path integral of the quantum fluctuations around a particles classical path. As with the Density Gradient formalism, the Effective Potential method has been shown to capture the quantum mechanical shift in threshold voltage, though, while describing well the shift in peak electron concentration away from the interface, it fails to capture accurately the shape of the electron distribution [113]. The Effective Potential approach also fails in capturing accurately tunneling effects [117].

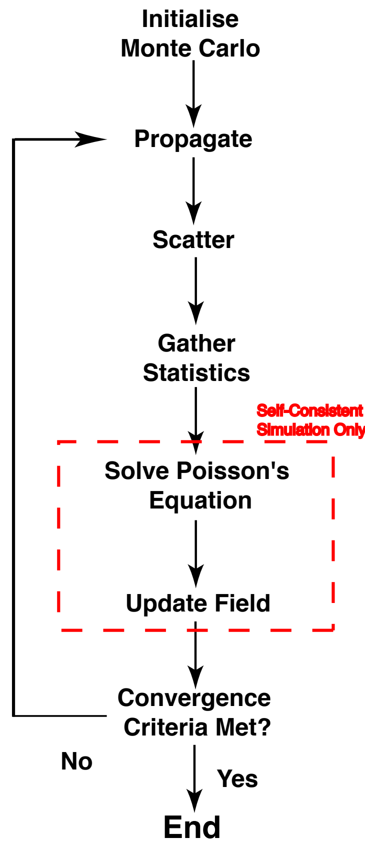

Figure 3.3: Simulation flow for Monte Carlo.

3.3 Monte Carlo

The Monte Carlo technique [106, 120, 121, 122, 123] offers a numerical solution to the BTE (3.1), combining classical (Newtonian) mechanics with quantum mechanical scattering rates. The method is based around the simulation of the trajectory of carriers in the presence of applied electric fields and random scattering events.

A typical simulation flow for the ensemble MC method is shown in Fig. 3.3. The simulator is initialised, populating the solution domain with carriers in a first guess which can be based on an analytical model, previous results, or

on the solution from another simulator, such as DD. Carriers are then moved under the influence of the field distribution in the device for a 'free flight' time generated via a random number. At the end of this free flight, a scattering mechanism is selected, again via the generation of a random number. These mechanisms replicate the correct physics via appropriate probability distributions and selection of post-scattering state. Fig. 3.4 shows a typical relationship between scattering mechanism and random number, generated via probabilities associated with each scattering process considered in the simulation. Fig. 3.5 shows the cumulative scattering rates for some of the scattering mechanisms used in our simulations, plotted versus energy. In addition to relevant, physically accurate processes, a 'self-scattering' (no change in the carriers state) is included to make the selection of the flight time easier as the scattering rate becomes constant over the energy range. Usually the self-scattering is the most frequently chosen scattering mechanism.

Once scattered, statistics are gathered and averaged, in terms of the individual attributes of the particles (velocity (3.12) and energy (3.13) - both for non-parabolic band approximations, with a non-parabolicity factor α), as well as ensemble averages such as current.

$$\mathbf{v} = \frac{\hbar \mathbf{k}}{m^*(1 + 2\alpha E(\mathbf{k}))} \tag{3.12}$$

$$E(\mathbf{k})(1 + \alpha E(\mathbf{k})) = \frac{\hbar^2 k^2}{2m^*} \tag{3.13}$$

The simulation can be 'frozen field' [124, 125, 126] or self-consistent [82, 127, 128]. In the 'frozen field' simulations the field distribution in the simulated device remains fixed, generated, for instance, using an initial DD simulation. In self-consistent simulations the new electron distribution at the end of each time step is used to solve Poisson's equation (3.2), and a new field distribution

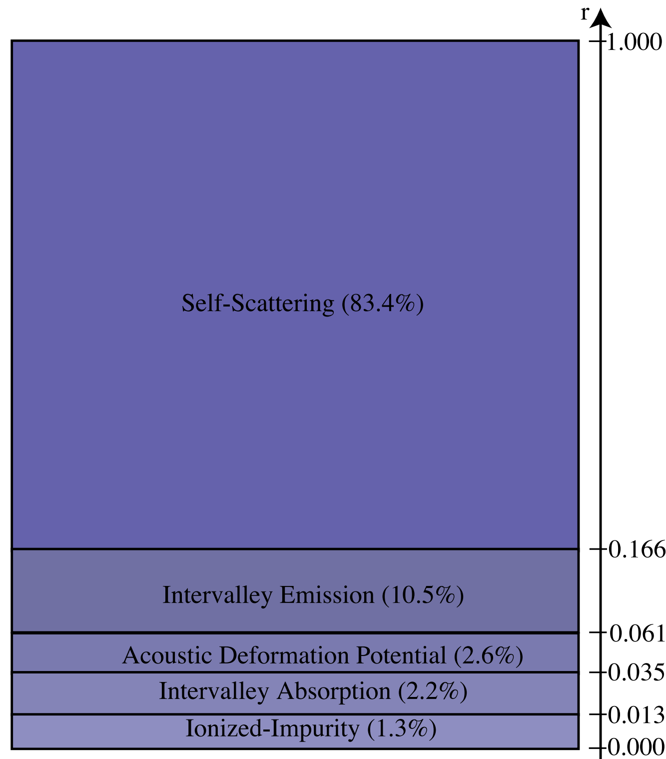

Figure 3.4: Relationship of scattering rates to random number, adapted from [106].

is calculated. The process is continued until statistics converge to a satisfactory accuracy.

The biggest detriment to computational speed is from solving Poisson's equation and the use of an efficient solver (such as SOR (Successive Over Relaxation) [88, 129], Bi-Conjugate Gradient-Stable (BiCGstab) [130] or Multigrid [131, 132] methods) is essential.

The band structure can be described in a succession of increasingly more complex models, starting with a simple, single band, parabolic/non-parabolic approach [120], which uses spherical or ellipsoidal bands, to a more detailed full

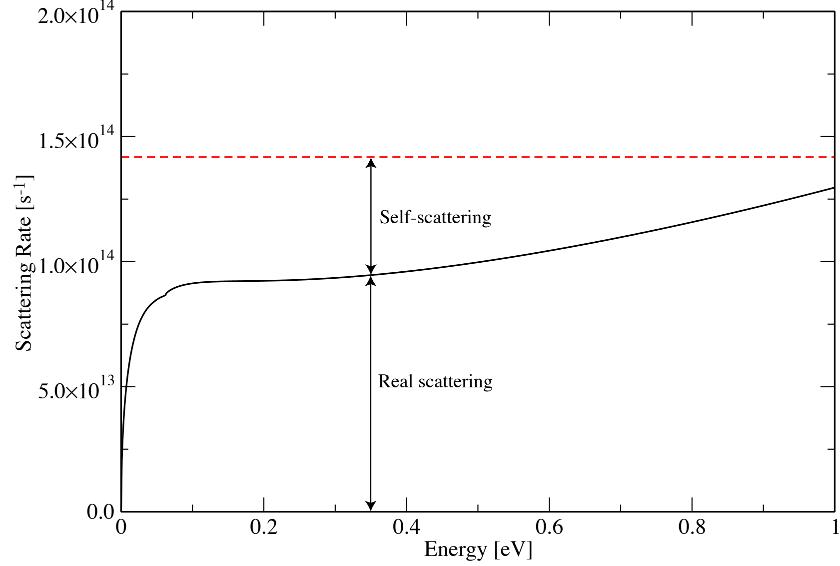

Figure 3.5: Cumulative scattering rates from the MC simulator used in this work, along with the total scattering rate taking into account self-scattering.

band structure based on the $\mathbf{k} \cdot \mathbf{p}$ [122], pseudopotential [125, 133, 134] or tight-binding [135, 136, 137] approaches. The increase in the complexity of the bandstructure description leads to a trade-off between accuracy and computational efficiency.

The clear benefit of MC over DD based simulations is its ability to handle transport in aggressively scaled devices where the number of scattering events is reduced leading to non-equilibrium and near-ballistic transport. The reduced number of scattering events leads to increasingly non-equilibrium transport with carrier heating [138], and non-local effects such as velocity overshoot becoming more and more relevant [16, 107]. As the MC approach is effectively a numerical solution to the BTE, rather than a simplification, it remains valid for small devices. Additionally, it is relatively easy and computationally efficient to extend MC into a 3D simulation domain.

3.3.1 Scattering in Monte Carlo

All of the transport phenomena described in Chapter 2 must be represented accurately in order to properly model UTB MOSFETs as they are scaled to t_{Si} below 10 nm. The inclusion of the corresponding scattering processes in simulations is therefore essential to capture the drive current magnitude and variations in addition to electrostatic effects which mainly influence the subthreshold current and the threshold voltage. In doing so a balance has to be achieved between computational efficiency and the physical accuracy.

The standard method for inclusion of scattering in MC has been described before, with scattering processes chosen via the selection of a random number compared to a table of scattering probabilities [121].

The probability of transition from state $\mathbf{k}$ to $\mathbf{k}'$ ($S(\mathbf{k}, \mathbf{k}')$) due to a perturbation Hamiltonian $H_{\mathbf{k}',\mathbf{k}}$ is given by Fermi's Golden rule (3.14), derived from the Schrödinger equation [106, 123]:

$$S(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} |H_{\mathbf{k}',\mathbf{k}}|^2 \delta[E(\mathbf{k}') - E(\mathbf{k})] \tag{3.14}$$

From this, scattering probabilities are calculated and stored for each mechanism. This method allows for scattering processes to be captured in detail as Hamiltonians can be calculated for each relevant scattering mechanism. This formalism has been used to calculate the scattering rates for Phonons (acoustic and optical) [120, 122], ionized impurities [106, 120] and interface roughness [139] scattering. An alternative method for including surface roughness scattering by weighted selection of either a specular or diffusive reflection [73] has been employed in [140]. This method has the drawback of requiring a varying ratio of specular to diffusive scattering depending on the bias applied to the device [141].

After scattering, a new final state is chosen, with the energy, the direction and the magnitude of the momentum (k-vector) of each scattered particle altered appropriately, based on the nature of the scattering mechanism [106]. The specific methods employed for these mechanisms can be found in various text books and papers [106, 121].

Another method for introducing scattering in Monte Carlo simulations is the so called '*ab initio*' scattering approach where the effect of the scattering potentials is captured through its effect on the real space trajectories rather than through probability-based scattering rates selected via random numbers.

It has previously been used to capture the influence of Coloumb scattering from a trapped charge or random ionized impurities [142, 143]. In this method, based on a particle-particle-particle-mesh (P^3M) algorithm [144], the short range Coulomb force acting on the real space trajectory of each particle can be evaluated using:

$$F_{SR}(\mathbf{r}) = \frac{Qr}{4\pi\varepsilon(r^2 + \frac{1}{2}r_c^2)^{3/2}} \tag{3.15}$$

Where r_c is a cut-off radius selected to prevent undesirable carrier heating [145] and r the distance to the charge. The long range interactions are included via the solution to Poisson's equation. This allows the accurate simulation of random dopant scattering induced current variability to be simulated [146].

This has proved to be a computationally efficient method that could capture not only the localised carrier-impurity interactions but the carrier-carrier scattering in small devices [146].

3.3.2 Quantum Corrections in Monte Carlo

MC simulations can effectively bridge the gap between the semiclassical (transport and scattering) and quantum (quantization and tunneling) regimes [141] through the introduction of quantum corrections. The transport phenomena detailed in the last chapter can all be adequately captured by quantum corrected MC down to very small dimensions ($t_{Si} < 5$ nm) [147, 148].

Hence, the use of quantum corrections within MC allows for computational efficiency to be maintained while the additional effects of size quantization and tunneling can be captured, thus extending the validity of the method. Various methods [149] based on Schrödinger equation [83], Wigner equation [150] and Effective Potentials [86] have been used.

In all cases, the general method involves the use of quantum corrected potential (ψ_q) [149] in the form:

$$\psi_q = \psi_c + \psi_{qc} \tag{3.16}$$

where ψ_c is the classical potential from the solution to Poisson's equation and ψ_{qc} is the quantum correction potential calculated using one of the methods described below. However, in contrast to DD simulations, instead of using the quantum corrected potential alongside the current continuity equation, in MC the corrected potential is used to calculate the force, F acting upon the particles during the free flight, where:

$$F = F_c + F_{qc} \tag{3.17}$$

$$F_c = -\nabla\psi_c \tag{3.18}$$

$$F_{qc} = -\nabla\psi_{qc} \tag{3.19}$$

This method of adding a correction term to existing semiclassical transport is considerably more efficient than solving the Poisson-Schrödinger equations which, while more accurate in capturing the full quantum picture, is usually restricted to 1D or 2D MC simulations [64, 139, 151, 152]. A more efficient approach to include Poisson-Schrödinger solution has been proposed in [153] for 2D MC simulations, where the Schrödinger equation is only solved periodically, and at the rest of the time steps, the eigenenergies are approximated using a perturbative approach.

A correction based on the Wigner distribution function has been proposed and employed in [154]. Starting with a modification to the BTE known as the Wigner transport equation (3.5), the correction term can be written as [84]:

$$\psi_{qc} = \frac{k_B T}{24}[\gamma^2(k-\overline{k})^2 - 3\gamma]\nabla\ln(n) \tag{3.20}$$

Where $\gamma = \hbar^2/(m^* k_B T)$ and $\overline{k}$ is the average momentum. As with the Density Gradient formalism, by averaging out the momentum term and assuming thermal equilibrium, a simplified quantum correction potential is derived from [84]:

$$\psi_{qc} = -\frac{\hbar^2}{12m^*}\nabla^2\ln(n) \tag{3.21}$$

This correction is added to ψ_c according to (3.16), and the field calculated according to (3.17) and (3.19). As it is based on the 2nd derivative of the carrier concentration, this approach is sensitive to noise, which is a clear drawback in MC simulations where the statistical carrier distribution is far from smooth. Therefore, the electron concentration has to be time averaged, and the potential spatially smoothed [155, 156]. So far, this method has been em-

ployed in 2D MC simulations [157], and has been shown to agree well with Poisson-Schrödinger [155] and NEGF results [158].

The above method was extended further in the so called effective conduction band edge (ECBE) method, considered a variation to the Wigner correction described above [85], or derived directly from the Bohm interpretation of the Schrödinger equation [109]. Here, the problem of calculating the driving force from higher derivatives of the electron distribution is avoided using an expression for $n \propto \exp(q\psi_q/k_BT)$. On this basis, the corrected potential can be written as:

$$\psi_q = \psi_c + \frac{\hbar^2}{4m^*rk_BT}[\nabla^2\psi_q - \frac{1}{2k_BT}(\nabla\psi_q)^2] \tag{3.22}$$

While (3.22) is less sensitive to noise than the previous methodology, it is strongly non-linear, making it difficult to solve. This correction has been successfully demonstrated in 2D MC simulations [159, 160] and extended to include the electron distribution across different valleys [161, 162, 163], as well as beyond the thermal equilibrium assumption [160].

The Effective Potential approach (3.9)-(3.11) mentioned above has also been successfully employed to capture confinement effects in 2D [164] and 3D [86] MC simulations. It has been used to study surface roughness scattering [165, 166] in DG MOSFETs. Here, selection of a weighted specular/diffusive reflection [73] is not appropriate as the extent to which carriers are repelled from the interface is overestimated by the Effective Potential method, so a scattering rate has to be employed to circumvent this issue and to allow for accurate simulation. Coulomb interactions with unintentional doping [62, 167] have also been studied with this method, and effectively quantifies the degradation due to dopant placement within the channel

A significant drawback comes through use of the fitting parameter, a_0 [168]. The underestimation of the electron concentration at the interface due to the

overestimation of the field, leads to a shift of the peak concentration away from the interface. Reducing this parameter to compensate for this can lead to a correction term smaller than expected and an undesirably large peak concentration when compared to a Poisson-Schrödinger solution. Extending to 3D requires varying values of the fitting parameter in the directions normal and parallel to the interface [167].

An alternative Effective Potential method has recently been proposed using a Pearson IV distribution in place of the Gaussian [169, 170]. The Pearson distribution can be tuned at various points depending on the distance of the carrier from the semiconductor-oxide interface in the vertical direction based on a total of four parameters compared to the single fitting parameter used for the Gaussian based Effective Potential. This allows a closer agreement with the electron distribution obtained from a Poisson-Schrödinger solution [170]. So far this has only been implemented and validated in 2D [169] and to the best of our knowledge has not been used in any detailed simulation studies.

While the inclusion of a Poisson-Schrödinger solution directly into MC simulation is computationally inefficient, a Schrödinger based correction can be employed via [50]:

$$\psi_q(z) = -k_B T \log(n_q(z)) - \psi_c(z) + V_0 \tag{3.23}$$

Here, $n_q(z)$ is the quantum density as taken from the solution of the Schrödinger equation [83], but represents only the shape of the distribution, rather than the actual values. In this case, the eigenvalue solver is only used in slices along the device in the quantization direction, and applied to calculate $n_q(z)$. This is done periodically to maintain self-consistency, and the result used to update the correction in (3.23), but not in conjunction with the Poisson equation which is solved using the MC electron distribution. Hence, the computational efficiency of a classical simulation is maintained. As only the shape of the quantum carrier density is used, the Fermi level is not required for this method and,

by use of different carrier temperature in each slice, it can be extended beyond the thermal equilibrium assumption [83].

Since this method uses the solution to the Schrödinger equation, good agreement with Poisson-Schrödinger solution is obtained. This methodology has been employed in both 2D [83, 171] and 3D MC simulations [172] (the correction remains 2D while transport is 3D), and captures size quantization accurately, but is not suited to tunneling effects [149].

Each of the described methods has its own strengths and weaknesses, both in terms of computational efficiency and accuracy. As a result no method has proved itself definitive, and choice of correction is based on suitability to purpose.

With reference to Table 3.1, a brief comparison between the different quantum correction approaches in MC simulations can be made. The Effective Potential and Schrödinger based versions have the benefits of being conveniently extendable to 3D simulations (at least in terms of transport), allowing for effects in the width direction to be taken into account (for instance Coloumb or surface roughness scattering) and neither of them are sensitive to the noise contained in the MC statistics. However, both are unable to effectively capture tunneling, and the Effective Potential approach overestimates the drop-off in electron concentration towards the oxide interface. The one significant difference between the two approaches is that the Effective Potential method requires a fitting parameter (a_0, or more specifically m^*) to properly match the solution from a quantum approach (such as Schrödinger-Poisson), whereas the Schrödinger based correction does not, as it uses a physical value for the effective mass m^* in its solution. An alternative implementation of the Effective Potential using a Pearson IV distribution in place of the Gaussian has recently been proposed, and circumvents the problems associated with the electron distribution drop off through the introduction of additional fitting parameters. At the time of

Effective Potential	3D, simple to implement, not sensitive to noise, fails to capture tunneling, overestimates drop off in the electron distribution towards the interface, uses fitting parameter
Wigner Correction	2D, sensitive to noise, mimicks tunneling, uses fitting parameter
ECBE	2D, less sensitive to noise, non-linear, mimicks tunneling, uses fitting parameter
Schrödinger based	Applicable to 3D, not sensitive to noise, accurate for quantization, does not capture tunneling, no fitting parameter

Table 3.1: Strengths and weaknesses of quantum correction methods in MC.

writing, this method is still in the validation stages, and has so far only been implemented in 2D.

The Wigner and ECBE corrections both share many common aspects, as they are variations on the same approach. Both have so far been extended only to 2D, both can mimic tunneling and both use a fitting parameter (m^*). The difference is in their sensitivity to statistical noise present in MC simulations. The Wigner corrections rely on the second derivative of the electron concentration and require smoothing to maintain computational efficiency. ECBE overcomes this problem via substitution, so is not directly dependent on the electron concentration, but at the cost of increased non-linearity, inserting its own problems in solution.

3.4 Quantum Transport

It has been suggested that semiclassical approaches can retain validity down to channel lengths of 10 nm [95], but, quantum effects, and in particular tunneling, will be playing an increasingly important role below this mark. To capture accurately such effects, it is necessary to include coherent transport self-consistency with the solution of Poisson's equation. For studying variability, such simulations have to be carried out of statistical ensembles of unique devices. One such method is the NEGF formalism [173, 174, 175], which directly captures quantization and tunneling.

The NEGF approach provides an open boundary solution to the Schrödinger equation. In this approach the retarded Green's Function, G^R, based on the relation shown in (3.24), is used to describe electron dynamics, by means of giving the response at one point in the system based on the excitation at another [176].

$$G^R(E) = [EI - H - \Sigma_1^R - \Sigma_2^R - \Sigma_S^R]^{-1} \tag{3.24}$$

Here, E is the energy, H is the Hamiltonian describing the system (for instance the effective mass Hamiltonian) and Σ^R is the self-energy used to describe the connection to the contacts (Σ_1^R and Σ_2^R) and the scattering mechanisms (Σ_S^R).

The electron density and current can be obtained from the diagonal of the lesser Green's function, $G^<$, that identifies how many states are occupied, and is defined by:

$$G^<(E) = G^R(E)\Sigma^<(E)[G^R(E)]^+ \tag{3.25}$$

Where $\Sigma^<$ is the inscattering function and $[G^R(E)]^+$ the transpose of $G^R(E)$. The greater Green's function, $G^>$, identifies how many states are empty and

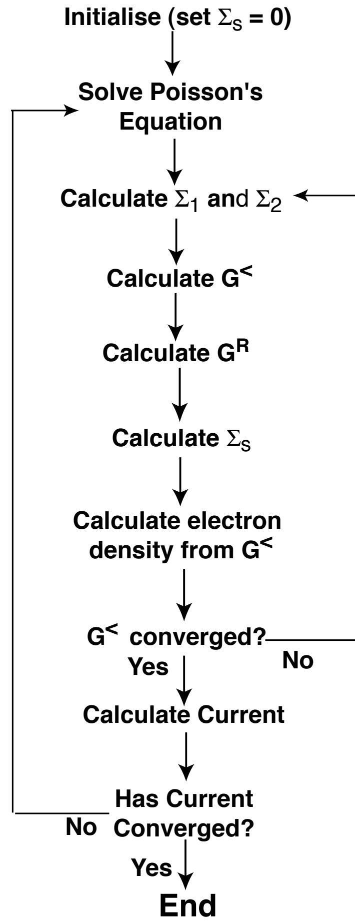

Figure 3.6: Simulation flow for NEGF.

the local density of available states can be obtained from the diagonal of the spectral function, $A = G^< + G^>$. The scattering functions are computed based on the desired process and approximation used. With $\Sigma^>$ the outscattering function, Σ_S^R is defined by:

$$\Sigma_S^R = \Sigma^< + \Sigma^> \tag{3.26}$$

Fig 3.6 outlines the flow diagram of a typical NEGF simulation, self consistently coupled to the solution of Poisson's equation.

The NEGF formalism includes quantum coherence effects that DD and MC simulations cannot properly capture including tunneling [176], and has been useful in calibrating and validating quantum corrected semiclassical models [110]. The greatest drawback comes from the computational cost associated with the formalism as large matrix transformations are required. This drastically increases as the complexity of the algorithms, and makes difficult the increase of the dimensionality of the simulation domain. These factors makes complete transport simulations difficult to achieve.

3.4.1 Scattering in NEGF

Ab initio elastic scattering from random discrete dopants, rough interfaces and body thickness variations are naturally included in the NEGF simulations through the associated Coulomb potential and boundary conditions at the interface. Inelastic Phonon scattering processes can be included via the self energy function [173]. However, the computational strain already imposed on the system increases yet further depending on the methodology and level of approximation employed.

A simple method of including scattering via the self energies is to use an approximation based on Büttiker probes [177]. These probes are placed at points along the device and are used to perturb the energy and momentum of carriers in a manner that allows for the cumulative effects of all scattering effects to be captured [178, 179]. Thus, scattering is included without significant additional computational costs as each probe is effectively treated as an additional contact with current kept at zero [173] with electrons removed and then re-injected with altered energy and momentum. This method, however, uses a significant level of simplification and approximation.

The rigorous inclusion of Phonon scattering, for example, via the self-consistent Born approximation [180, 181], leads to substantial computational costs, as

large matrix inversions are required. Here, $\Sigma^<$ and $G^<$ for a given energy must be calculated self-consistently [173, 176], the nature of the relationship between the two described by the coupling constant D. This can be done in the case of elastic and inelastic scattering as shown in (3.27) and (3.28) respectively.

$$\Sigma^<(E) = DG^<(E) \tag{3.27}$$

$$\Sigma^<(E) = D_{ab}G^<(E - \hbar\omega) + D_{em}G^<(E + \hbar\omega) \tag{3.28}$$

In (3.28), the two terms on the right hand side deal with the absorption and emission of a Phonon respectively. Therefore, to prevent the computational costs becoming excessive, either the dimensions of the simulation or the set of scattering mechanisms considered are restricted, making it difficult to obtain a full transport picture. Hence a lot of work carried out using NEGF simulations is focused on ballistic transport in MOSFETs where scattering is not an issue [90], and is restricted to 1D [182] or 2D [175] simulation domains.

Ab initio methods have been implemented to study the impact of unintentional dopants and interface roughness in UTB DG MOSFETs.

The use of a NEGF simulation method restricted only to 2D leads to difficulties in capturing scattering from stray impurities. In this case, current flow in the width direction is not considered and the impact of individual charges is overestimated [183]. Full scale 3D simulation is required to properly capture the effect [141].

Scattering from interface roughness induced body thickness fluctuations, as described in the previous chapter, has been studied using 1D [184] and 2D [185, 186] NEGF simulators, both self-consistently coupled to Poisson's equation. In the 1D case, a DG device with a nominal value of $t_{Si} = 3$ nm was simulated, and a significant degradation in I_{on} due to body thickness variations

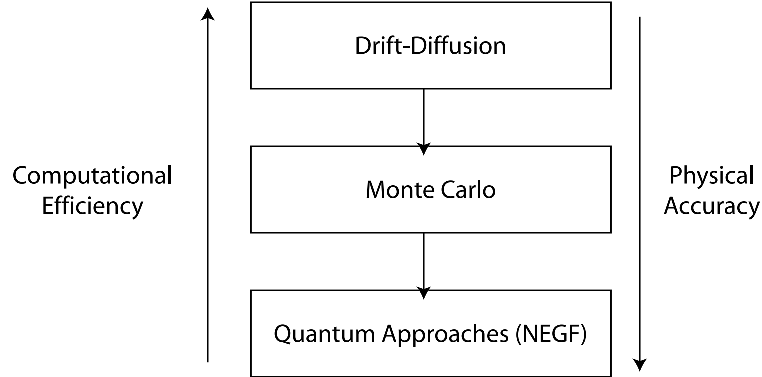

Figure 3.7: Hierarchy of simulation methodologys.

demonstrated. In the 2D study, a DG MOSFET was simulated, with $t_{Si} = 2$ nm, and the impact of the varying thicknesses on current flow has been successfully captured. However, in both cases, the restriction in dimensionality leads to a failure in realistically capturing the impact of interface roughness patterns, which are two dimensional in nature. While the 2D simulations capture the confinement scattering, they cannot describe the current percolation in the width direction. This study does highlight the usefulness of the NEGF formalism in capturing current variability due to the unique body thickness pattern in UTB nano CMOS transistors.

For more realistic simulations of the above effects it is necessary to extend the simulation domain to three dimensions but, the computational strain associated with NEGF simulations becomes prohibitive. As a result, implementation of the above *ab initio* scattering methods into 3D has so far been restricted to the simulation of a $2 \times 2 \times 6$ nm nanowire [187].

3.5 Summary

Fig. 3.7 highlights the trade-off between computational efficiency and physical accuracy in the choice of simulation methodology.

As the focus of this research is the impact of interface roughness and the corresponding body thickness variation induced scattering on transport and current variability in UTB MOSFETs, it is useful to compare and contrast the suitability of the simulation methods described in this chapter to this purpose.

DD is efficient and stable simulation approach, and can include quantum effects in a simple but effective manner, leading to a useful simulation tool. However, it cannot capture non-equlibrium effects in small devices and has difficulties in capturing the transport variability associated with the difference in the body thickness pattern from device to device. The latter limits its usefulness in this study.

NEGF based simulations offer accurate description of quantum transport in a manner not available from the other two simulation methods. It can handle in an *ab initio* manner the scattering introduced by the body thickness variations. However, to accurately capture the impact of this mechanism, 3D simulations are necessary, which are still computationally prohibitive in NEGF simulations of devices of relevant CMOS size. Their use in statistical simulations, central to our research, is out of the question. The best hope is that NEGF simulations can provide a basis for calibration and validation of the DD and MC results.

MC simulations capture non-equilibrium effects, can include quantum corrections while maintaining computational efficiency and accuracy and can handle *ab initio* scattering in statistical simulations of variability. In quantum corrected MC, the quantum transport is treated phenomenologically, but scattering is treated in detail.

The selection of simulation technique is based on the importance of the investigated phenomena and the trade off between the accuracy and the efficiency by which this phenomena is represented in each one of the simulation techniques. While all the techniques considered in this section offer advantages and disadvantages in terms of both the effects captured and the computational efficiency, it was decided that it is most appropriate to use a MC technique in this study, as it offers the best trade off in describing the body thickness fluctuations related variability in nanometer scale UTB DG or SOI devices with an acceptable level of computational efficiency.

Chapter 4

Simulator Development

4.1 Introduction

In order to study the impact of scattering from interface roughness induced body thickness fluctuations on the variability of ultra thin body silicon on insulator (UTB SOI) MOSFETs, an appropriate 3D simulator capable of dealing with non-equilibrium transport and transport variations from device to device had to be developed. It was necessary to use a 3D simulation domain in order to correctly account for percolation across the width of a device, that would be neglected in a 2D simulation. The computational efficiency becomes a very important factor in such 3D simulations, therefore a 3D Monte Carlo (MC) simulator has been developed which includes scattering from the interface roughness and body thickness variations both captured via the real space trajectories of the MC particles.

In this chapter, the development of this simulator is detailed. Firstly, the Glasgow 'Atomistic' Drift-Diffusion (DD) Simulator, which has been used in conjunction with the MC simulator, including the technique for generation of the interface roughness pattern, is described. The emphasis is on the modifica-

tions made to this simulator through the course of this project. This is followed by the description of the MC module starting with a simple, 'frozen field' approximation, and moving on to self-consistency with Poisson's equation and the inclusion of quantum corrections. Additional considerations concerning the inclusion of interface roughness and the proper implementation of contacts are also described at the end of the chapter, along with a discussion of the selection of optimal time steps and mesh spacings for MC simulations.

4.2 Drift Diffusion

The Glasgow 'Atomistic' 3D DD simulator [188] has been developed and maintained over several years by different members of the Device Modelling Group. In this project, the focus was on using it to provide input for the MC simulations. Additionally it was used as a benchmark for comparison and validation of results obtained from the MC module. Alterations were made to the existing 'Atomistic' code to allow extraction of the necessary data from simulations for the purposes of this project. This section gives a brief overview of the operation of the simulator, focusing on elements important to the MC module.

The 'Atomistic' simulator allows for a variety of sources of variability to be studied, such as line edge roughness (LER), random discrete ionized dopants and interface roughness [92, 93] in a variety of structures ranging from conventional MOSFET architectures, to SOI [115] and double gate (DG) [189] and nanowires. It is most suited to sub threshold regimes where the current is dominated by the device electrostatic and is less affected by transport variations. As a result, the development of a 3D MC simulation methodology is essential to enable the study of non-equilibrium transport and on-current (I_{on}) variations resulting from scattering sources, in this case the effects of non uniform oxide and silicon layers.

The DD simulator follows the basic methodology described in the preceding chapter. All equations are discretised using a finite box method onto a mesh covering the solution domain which is based on dimensions supplied by an input file. A uniform mesh has been employed in simulations used to initialise the MC module for reasons which will be detailed in the course of the this chapter. A Red-Black Successive Over-Relaxation (SOR) scheme is used for the solution of Poisson's equation, and a Bi-Conjugate Gradient Stabilized (Bi-CGSTAB) solver is used for the continuity equations. Quantum corrections are implemented via the Density Gradient formalism:

$$\phi_n - \psi + \frac{k_B T}{q} \ln\left(\frac{n}{n_i}\right) = 2b_n\left(\frac{\nabla^2\sqrt{n}}{\sqrt{n}}\right) \tag{4.1}$$

$$b_n = \frac{\hbar^2}{12qm^*} \tag{4.2}$$

Where ϕ_n is the quasi-Fermi level, ψ the potential, k_B is Boltzmann's constant, T the temperature and q the electron charge. A discretization and solution scheme similar to the one used for Poisson's equation is employed. This is solved self-consistently with Poisson's equation and the current continuity equation iteratively in order to find the quantum corrected charge density. With this, a new quantum corrected potential can be evaluated and the procedure iterated until the current converges.

At the Si/SiO_2 interfaces, the electron density goes towards zero due to a large potential barrier being present [51]. Initially, this was taken into account in the solver by fixing the interface concentration to an arbitrarily small value, assuming an effectively infinite barrier. A better solution was proposed in [190], taking into account that the barrier is actually finite and that there is penetration of the electrons wavefunction into the oxide. Therefore, the boundary

	Δ [nm]	L_c [nm]	Auto-Correlation Function
Goodnick *et al.* [191]	0.48	1.3	Exponential
Goodnick *et al.* [191]	0.35	1.5	Gaussian
Yoshinobu *et al.* [192]	0.3	15	Exponential
Yamakawa *et al.* [139]	0.178	2.2	Exponential
Vasileska *et al.* [193]	0.3	1.5	Exponential
Gamiz *et al.* [74]	0.5	1.5	Exponential
Winstead *et al.* [83]	0.178	2.2	Exponential
Esseni *et al.* [77, 194]	0.51	1.0	Gaussian

Table 4.1: RMS height, correlation length and auto-correlation used (or shown to be preferable) in both simulation and experimental studies of semiconductor/oxide interface roughness. In the case of Goodnick, both auto-correlation functions were employed without a strong preference being stated.

conditions at the interface between oxide and semiconductor has been modified to:

$$n \cdot b_n \nabla \sqrt{n} = -\frac{b_{ox}\sqrt{n}}{x_p} \tag{4.3}$$

Where $b_{ox} = \hbar^2/(12qm^*_{ox})$ and $x_p = \hbar/(\sqrt{2m^*_{ox}\Phi_B})$ is the characteristic penetration depth into the oxide [190] with Φ_B the silicon to insulator potential barrier. This makes the carrier density at the interface an unknown variable allowing for a smooth transition of the carrier density towards the interface that is representative of the electron penetration into the oxide, rather than the sudden discontinuity associated with the assumption of zero electron charge at the interface.

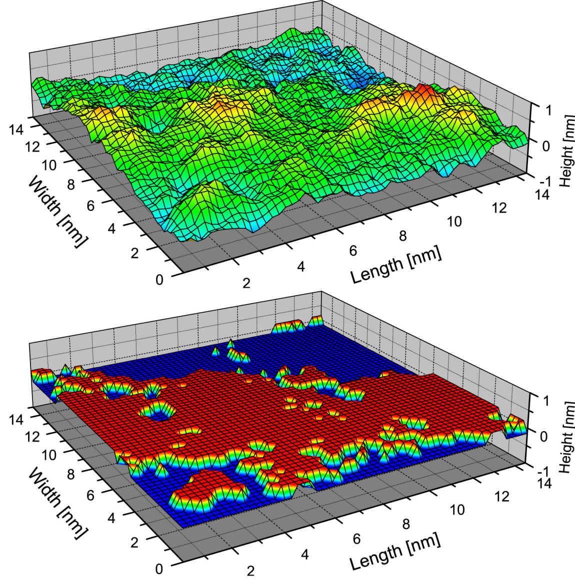

Figure 4.1: Generated roughness pattern (above) and the corresponding digitised version (below) from [115]

4.2.1 Interface Roughness Generation

The DD simulator has a subroutine dealing with the generation of rough surfaces at the semiconductor/oxide interface. The roughness is described statistically as a fluctuation in oxide thickness, $\Delta(r)$, with a RMS amplitude, Δ, and correlation length, L_c, based on the methodology described in [92, 115, 195]. Values of $\Delta = 0.3$ nm and $L_c = 1.8$ nm were used in the simulations carried out in this project (these values are representative of a range of experimentally measured values [139, 192, 195]), along with an exponential auto-correlation function (4.4) which has been shown to be more appropriate than a Gaussian equivalent for representing surface roughness scattering in MC simulations [139, 191, 192].

$$ACF(r) = \Delta^2 \exp\left(-\frac{r}{L_c}\right) \tag{4.4}$$

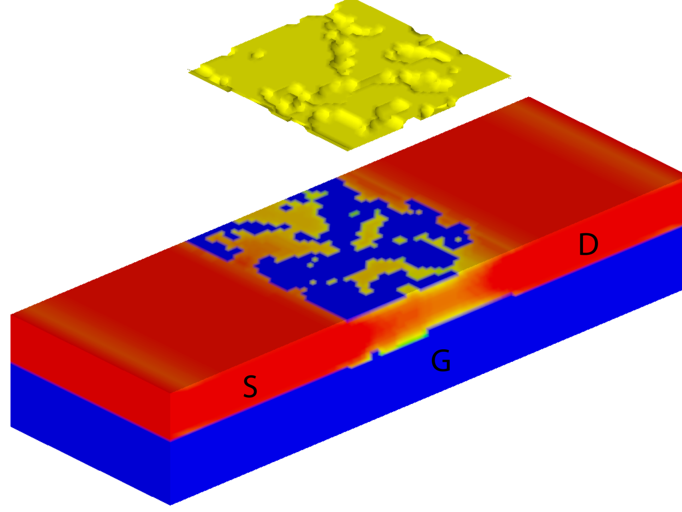

Figure 4.2: Impact of the inclusion of interface roughness on electron distribution in a 10 × 10 nm double gate device, modelled using DD, showing confinement of carriers along the channel. The source (S), drain (D) and bottom gate (G) are shown, with the top gate removed to show impact of the roughness pattern at the top of the channel.

The values of Δ and L_c are confirmed by a comparison of simulation and experimental studies of the interface patterns shown in Table 4.1. A power spectrum is obtained via a 2D Fourier transform of the correlation function. The interface can then be reconstructed using an $N \times N$ matrix, the magnitude of each element following the power spectrum with the phase chosen at random. The roughness produced is digitized to 0.15 nm steps above and below the position of the smooth interface to match the spacing of a silicon monolayer (see Fig. 4.1). In the vertical direction (top gate to bottom gate), the simulator mesh spacing is then chosen to match the 0.15 nm steps and a uniform mesh is used to avoid self-force in self-consistent MC simulations [196].

The variations in the thickness of the oxide, and consequently the thickness of the silicon body, have an effect on the electrostatics and the quantum confinement in the device, shifting accordingly the threshold voltage which is

well captured in the Density Gradient corrected DD simulations. Additionally, the non-uniformity of the silicon thickness along the channel produces an additional scattering potential due to variations in carrier confinement (see Fig. 4.2) as described in a previous chapter. This source of scattering is considered a significant source of transport degradation in UTB SOI and DG MOSFETs at the limits of scaling. However, the 3D DD simulator, whilst well able to predictively describe electrostatic effects, captures transport through a parameterized electron mobility model with little predictive power at these scales. Hence a 3D MC simulator is necessary in order to analyse the transport degradation and current variability due to increased scattering from quantum confinement variations.

4.3 Monte Carlo

The main focus of this project is the development of a 3D MC simulator, starting with a classical, 'frozen field' approximation, then adding self-consistency and quantum corrections. An *ab initio* implementation of scattering associated with interface roughness and the resulting body thickness variations in UTB DG and SOI MOSFETs was also developed in order to study current variability that would not be fully taken into account using DD simulations alone. Other relevant scattering mechanisms for phonon (acoustic and optical) and ionized impurity scattering are included via scattering rates calculated as described in the Chapter 3.

In the MC module, the charge is assigned to the mesh by virtue of the 'Cloud In Cell' scheme [144], where the charge density is interpolated to the eight grid points of the cell a carrier is currently in, based on its position relative to each corner. The equations of motion are integrated via the Velocity Verlet algorithm [197] and a spherical, non-parabolic band approximation is used [120].

4.3.1 'Frozen Field' Approximation

In its initial form, the MC simulator employed a 'frozen field' approximation [124, 125, 126]. The field was never updated to reflect the constantly changing carrier distribution, so there was no link between the dynamics of the carriers and the electric field. The benefit of the 'frozen field' approximation comes in its speed, with simulations taking as little as a matter of hours to complete. Therefore, statistical studies based on large ensembles of uniquely different devices may be easily and quickly carried out.

A 'frozen field' simulation consists of the following procedure. First, a DD simulation is run and the solution passed to the MC module in order to initialise the particles, and to calculate the driving force applied to them. The description of the mesh spacing in each direction, along with details of positions of specific architectural detail (start and end of gate and oxide regions for instance), applied bias (V_G, V_D), the electron, hole, donor and acceptor concentrations, the classical and (if necessary) quantum potentials are all passed from the DD to the MC simulator. While the majority of these parameters are required for MC simulator to function, some (generally the details concerning the structure and bias) are desirable only for record keeping, as they are written to a file at the beginning of the simulation. Additionally, as described previously, the interface roughness patterns are generated in the DD simulator, then applied in the MC module, the methodology for which is described later in this chapter. The MC simulator is capable of handling non-uniformity in the mesh spacing which, while acceptable for 'frozen field' solutions, can introduce self-force issues [196] in self-consistent simulations.

To begin with, the two simulators were completely separate entities, the DD simulator writing out a file with the relevant information that the MC module then reads in. To make the process more flexible and self contained, the two simulators were coupled via a very simple interface subroutine that is called once the DD simulation was completed, based on a flag within the input file

which is set to indicate that a MC simulation should be run using the output from the DD simulation.

This method allows for a single executable to be compiled and run, without the need for two separate simulations to be carried out independently. This was beneficial for simulating a variety of devices that were different in terms of, for instance, structure or bias. For testing purposes it was frequently useful to simulate identical devices, but using different MC parameters so the previous file based version was more applicable. Currently, the two sections still maintain individual input files, the DD one describing the mesh, applied bias and doping, the MC file containing details concerning time steps and number of particles. For further ease of use, these files could be combined into one single input file.

Further modifications concerned the languages the codes were written in. The DD code, having been developed over a longer period of time by various members of the Device Modelling Group, is coded in Fortran 77, while the MC module is in Fortran 90. This led to difficulties bringing them together. Converting the DD simulator to Fortran 90 would allow for arrays and solvers to be shared, and for memory to be dynamically allocated, lowering the overall memory footprint of the code and better using computing resources. This task remained outwith the scope of this project, which was to develop the MC part of the simulation process.

The main drawback of the 'frozen field' approximation comes from the fact that the method is limited to low field simulations, hence all simulations have to be carried out at low drain bias ($V_D < 50$ mV). This retains the validity of the 'frozen field' approximation by limiting the non-equilibrium effects, minimizing the importance of the coupling of the field to the current, and thus allowing comparison to DD simulations. Therefore, to capture a wider range of effects up to higher drain bias, the field would have to be updated self-consistently with the carrier distribution during the course of the simulations.

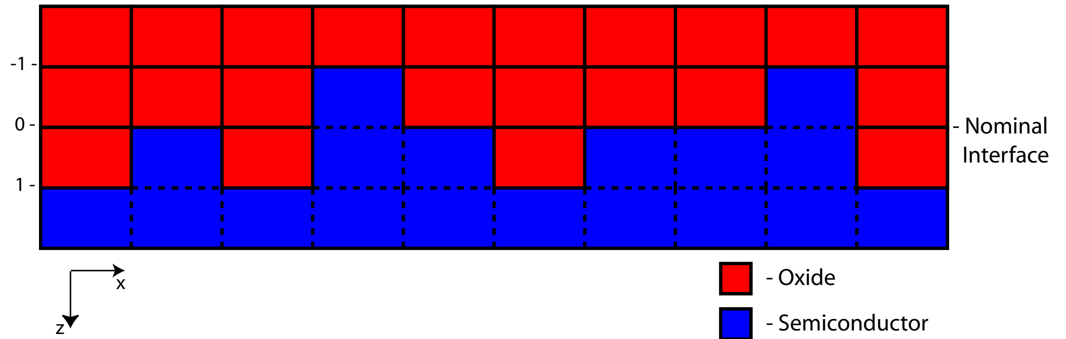

Figure 4.3: Digitization of interface roughness patterns in MC. As shown, the pattern is used to adjust the reflective boundaries and is applied consistent with the mesh used. This is carried out across the x-y plane for the entire depth of the device, running the length of the channel.

4.3.2 Interface Roughness Implementation

The implementation of an effective method to capture the scattering from surface roughness and body thickness fluctuations is central to this work. Two methods were developed to include the digitized interface roughness patterns generated by the DD simulator in the MC module.

The first, rather crude, method involved writing an array from the subroutine in the DD simulator that generated the pattern storing for each point in the x-y plane, depending on the value for the roughness at that point, either -1, 0 or 1 indicating its position relative to a nominal interface (see Fig. 4.3). Such an array for each interface is written by the DD simulator and then read into the MC module (either from a file, or via the interface) and the reflective boundaries at each semiconductor/oxide interface are altered accordingly. The roughness pattern extends for the length of the channel in the x-direction (not in the source or drain regions), and for the entire width of the device in the y-direction.

The second, more refined method uses an array created by the DD simulator that describes the material at each node. By passing the array from the DD to the MC module, the reflective boundaries could be altered in a more elegant manner, but still achieving the same outcome. In this way it was no longer necessary to explicitly state in the input file of the simulator that interface roughness was to be included in both the DD and the MC simulators as everything is defined at the start of the DD simulation, leaving less scope for the possibility of inconsistency. Both implementations allowed for an *ab initio* method of surface roughness scattering to be included via the real space interactions with the varying oxide roughness patterns.

Previously, scattering from body thickness variation has been included in MC simulations via a scattering rate [77], calculated to reproduce the specific mobility dependence on silicon thickness due to this mechanism, as discussed in Chapter 2. In this work, this mechanism is captured in an *ab initio* fashion via the real space trajectories of the particles by inclusion of an effective quantum scattering potential that is obtained from the silicon body thickness fluctuations as also described in Chapter 2. This is achieved by calculating the driving force acting on the particles by using a quantum corrected potential that takes into account the necessary changes in confinement and resulting shift in the ground state as the thickness of the channel varies. This makes the calculation of additional scattering rates and resulting final states unnecessary, while still taking into account the effects of this mechanism. The inclusion of quantum corrections in MC is dealt with later in this chapter.

4.3.3 Self-Consistent Monte Carlo

At higher V_D, the current and field are more strongly coupled and the potential must be regularly updated to reflect changes in the carrier distribution. The ability of MC simulations to capture non-equilibrium transport and the local variations in electron velocity that DD cannot, leads to a difference in the

resulting terminal currents between DD and MC. At low V_D, as simulated using the previous 'frozen field' approach, the non-equlibrium effects are limited, and similar results are recovered from both DD and MC simulations, but an increase in V_D will lead to different values being observed between the two simulation methodologies.

To implement self-consistency in MC [82, 127, 128], the classical potential, ψ_c, has to be updated at the end of each time step based on the new carrier distribution, n_{mc}, using Poisson's equation:

$$\nabla \cdot (\varepsilon \nabla \psi_c) = -q(N_D - n_{mc} + p - N_A) \tag{4.5}$$

Where ε is the permattivity for the particular medium and N_D, N_A and p the donor, acceptor and hole concentrations respectively. To solve this, two methods were considered, SOR [144] and Bi-CGSTAB [130]. Both were tested, and gave different results in terms of speed of solution, though they produced the same classical potential. Table 4.2 shows a comparison between them, indicating a considerable speed up when the Bi-CGSTAB solver is used, which is on average more than three times faster than SOR, and occupies a smaller percentage of the simulation time. Putting this into perspective, for a MC simulation of 200 000 times steps, using the Bi-CGSTAB solver the total simulation time would be less than four days, for the SOR solver it would be almost two weeks.

Additionally, the length of time step employed between Poisson solutions (Δt) should fit the relation $\omega_p \Delta t < 0.2$, where ω_p is the plasma frequency given by [198, 199]:

$$\omega_p = \sqrt{\frac{nq^2}{\varepsilon m^*}} \tag{4.6}$$

	Average Time per Solution	Percentage of Total Time
SOR	5.41 secs	97%
Bi-CGSTAB	1.52 secs	90%

Table 4.2: Comparison of average time of solution for the linear Poisson solution using Bi-CGSTAB and SOR for a tolerance of 10^{-6}. Also shown is the average percentage of the total simulation time spent solving Poisson's equation.

With m^* being the effective mass of the carrier. This maintains numerical stability as it prevents the generation of artificial plasma oscillations, which can be especially problematic in highly doped regions, such as the contact source and drain regions [198]. Equation 4.6 gives a time step of around 0.1 fs for $n = 2 \times 10^{20}$ cm^{-3}, which is typical for the source and drain regions of the devices simulated in this study. However, in the case of quantum corrected simulations, where the field towards the interfaces is higher than in classical simulations, an even smaller time step was required, usually around 0.01 fs, to avoid undesirable carrier heating. An even shorter time step (0.005 fs) was neceesary when body thickness fluctuations were introduced, due to localised regions where the potential varies rapidly, leading to higher fields and the possibility of even more artificial carrier heating. This was also a necessary consideration in 'frozen field' simulations. The unfortunate consequence of shortening the time step, is an increase in the number of total simulation time steps required for a complete simulation.

4.3.4 Quantum Corrections

As mentioned previously, inclusion of quantum corrections in the MC simulator is vital to this work. Methods to achieve this, based on the Density Gradient formalism, were introduced in the MC simulator in a succession of

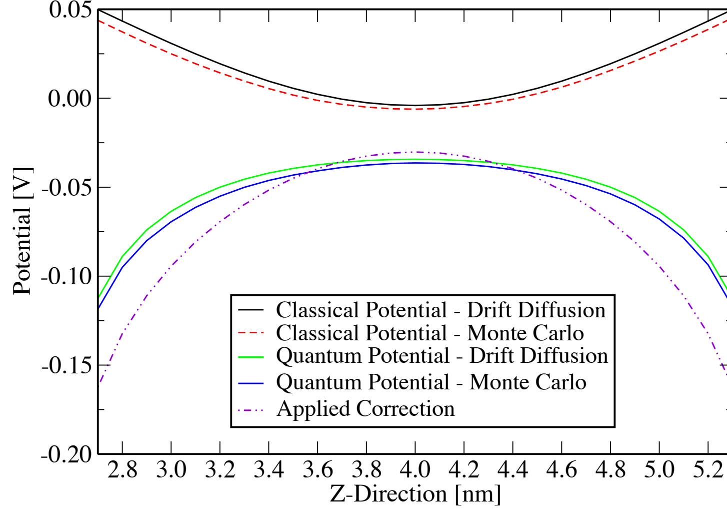

Figure 4.4: Classical and quantum potential along with the applied correction in the vertical direction (top to bottom gate) of a DG MOSFET (Device 2 in Table 5.1 in Chapter 5 with t_{Si} = 2.6 nm).

more complex and effective revisions that continually broadened the scope of simulations.

The first implementation of the Density Gradient quantum corrections was in the 'frozen field' approach, where the quantum corrected potential taken from a DD simulation with Density Gradient corrections is used to calculate the driving force applied to the particles. Additionally, the resulting quantum distribution of particles from the DD simulation is used to initialise the carriers at the beginning of the MC simulation. The field is not updated during the course of the simulation, so there is no self-consistent coupling of the MC carrier distribution to either the classical potential nor the quantum corrections.

The second update involves the use of a frozen quantum correction applied to a classically self consistent MC simulator. The self consistency is implemented as described previously, solving Poisson's equation at each time step, and again the quantum correction, ψ_{qc}, is calculated from an initial DD simulation by

taking the difference between the quantum corrected potential, ψ_q, and the corresponding classical potential at each point:

$$\psi_{qc} = \psi_q - \psi_c \tag{4.7}$$

This is saved, and after each solution to Poisson's equation (4.5), the correction is applied to the resulting classical potential:

$$\psi_q = \psi_c + \psi_{qc} \tag{4.8}$$

This procedure is illustrated in Fig. 4.4, showing the relationship between the classical and quantum potentials along with the applied correction in the vertical direction. This quantum corrected potential is then used to update the applied electrical field:

$$F_q = -\nabla\psi_q \tag{4.9}$$

Both these methods have the drawback of being restricted to low drain bias simulations, as they both have non-self-consistent elements to them. In the case of the 'frozen field' version, there is no connection between the current and the field. In the classically self-consistent version with a frozen quantum correction, the field is recalculated based on carrier dynamics, which does allow for higher drain bias to be considered, but the quantum correction is never updated to reflect changes in the system. As can be seen from (4.1), the quantum corrected charge density from Density Gradient is dependent on the quasi-Fermi level and classical potential from the solution of Poisson's equation, hence if these deviate significantly from their initial values (for instance due to non-equilibrium transport), there would be a resulting change in the quantum density, hence the quantum correction and electric field would be inconsistent with this new carrier distribution.

Ultimately, a code which updates the Density Gradient solution periodically would increase the range of validity of the MC simulator. For instance, this opens up the possibility of time dependent device simulations. To this end, a Density Gradient solver was introduced directly into the MC code. The implementation presented here is more in the spirit of the Schrödinger based corrections presented in [83], than the Wigner based corrections of [155].

The simulation is initialised as before, with the quantum correction, quantum potential and corrected field calculated using (4.7), (4.8) and (4.9) respectively and the simulation progresses initially as the classically self consistent version with a frozen correction. After the initial transient period, the values for the classical and quantum potential and the electron density are time averaged in order to smooth the numerical noise present in the MC simulator. After such averaging over a selected period of time, $\langle \psi_c \rangle_t$ and $\langle n_{mc} \rangle_t$ are used to update the quasi-Fermi level, ϕ_n (4.10), which itself is also time averaged.

$$\phi_n = \langle \psi_q \rangle_t - \frac{k_B T}{q} \ln \left(\frac{\langle n_{mc} \rangle_t}{n_i} \right) \tag{4.10}$$

Then, periodically, the quantum density, n_q, is obtained using (4.11), based on the values of $\langle \phi_n \rangle_t$ and $\langle \psi_c \rangle_t$.

$$\langle \phi_n \rangle_t - \langle \psi_c \rangle_t + \frac{k_B T}{q} \ln \left(\frac{n_q}{n_i} \right) = 2 b_n \left(\frac{\nabla^2 \sqrt{n_q}}{\sqrt{n_q}} \right) \tag{4.11}$$

Here, b_n is defined as before (4.2). With the new n_q, using (4.12) the new value for ψ_{qc} is calculated, and this correction is used until the next update of the Density Gradient solution.

$$\psi_{qc} = \langle \phi_n \rangle_t + \frac{k_B T}{q} \ln \left(\frac{n_q}{n_i} \right) - \langle \psi_c \rangle_t \tag{4.12}$$

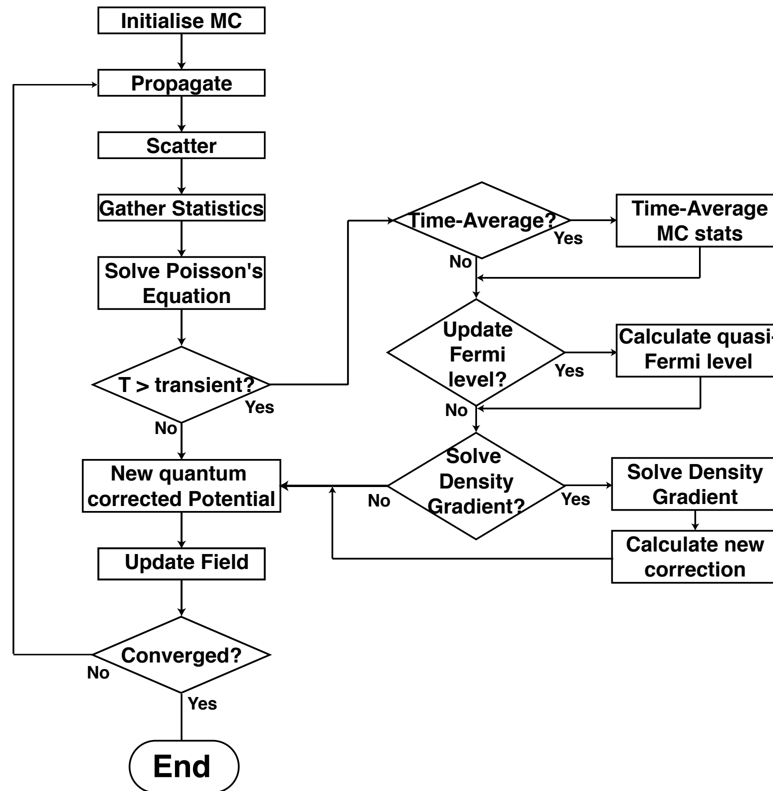

Figure 4.5: Flowchart of MC simulation with self consistent Density Gradient quantum corrections.

The solver uses a Red-Black SOR scheme, and the discretization is based on the finite box method, as employed for Poisson's equation. A flowchart for a MC simulation using this methodology for quantum correction is shown in Fig. 4.5, indicating the additional computational steps required to include the necessary calculations and time averaging. Fig. 4.6 shows how the simulation evolves starting with an initial transient period, followed by time averaging, then the periodic update of the quantum correction.

The frequency of solution of the Density Gradient equation in order to recalculate the quantum correction is still under investigation. In [50], the update is carried out every 100 fs using a Schrödinger based quantum correction. That

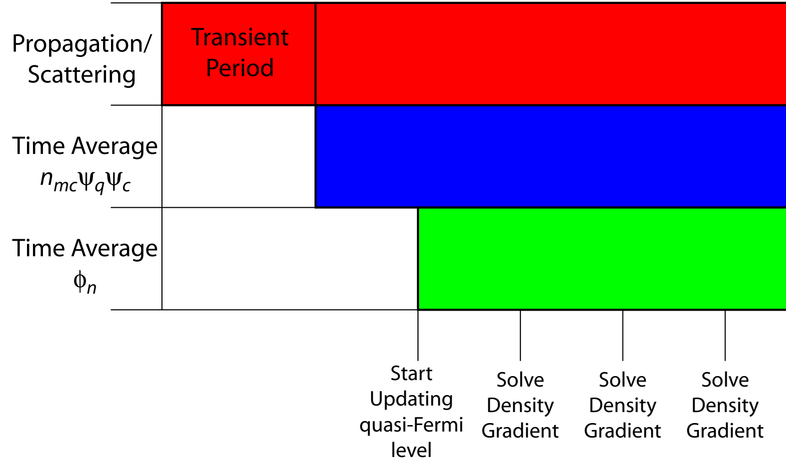

Figure 4.6: Timeline of fully self consistent MC simulation with Density Gradient quantum corrections, showing the order that processes begin within the simulation.

methodology is similar in spirit to that developed in this work, so would seem a reasonable place to start, and has thus far proved effective. Fig. 4.7 shows a comparison between the carrier concentration and quantum corrected potential from DD and the MC simulator using this method for quantum corrections. A good agreement is shown between the two methodologies at low V_D validating the use of this method of quantum correction in the 3D MC simulator developed as part of this work. Further testing may lead to refinement of the frequency that the Density Gradient equation is solved, perhaps based on applied bias or device dimensions.

Time averaging is conducted in the period between solutions, then reset immediately after the correction has been updated. Therefore, the new correction reflects the changes within the system in the intervening period. Another area of consideration is the frequency that values are gathered for time averaging, perhaps every time step (an interval of around 0.01 fs) or every 10 time steps (an interval of around 0.1 fs). Again, this may be dependent on the nature of the simulation, in terms of the applied bias, the size and architecture of the

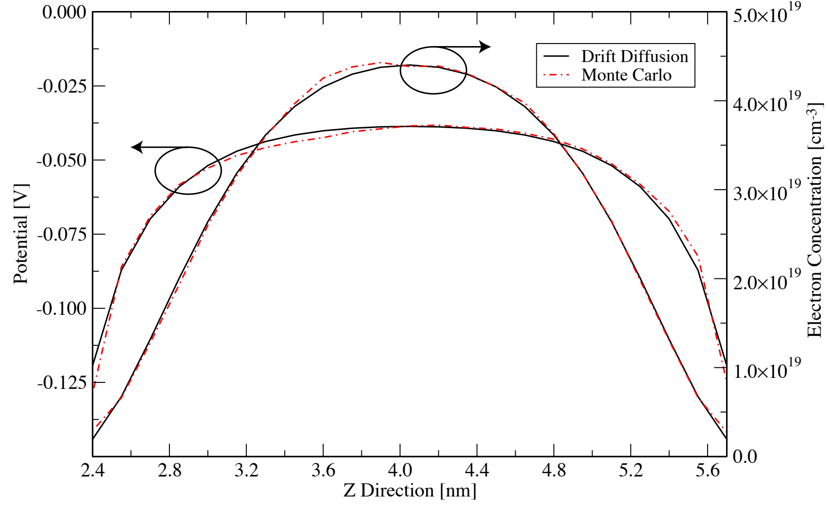

Figure 4.7: Quantum corrected potential and electron concentration in a DG MOSFET (Device 2 in Table 5.1 in Chapter 5 with t_{Si} = 3.3 nm) at V_D = 0.01 V and V_G = 0.8 V. Results from DD and MC simulations, both using self-consistent Density Gradient quantum corrections are presented, showing a good agreement between the two. Slight asymmetry due to noise in MC statistics.

device and the inclusion of a non-uniform silicon channel thickness. Therefore, further investigation is required into this matter. A further assumption made in the implementation of the Density Gradient approach is that the carriers are in thermal equilibrium with the lattice. This is a reasonable assumption only at low drain bias, but, as the bias is increased, and carriers gain energy in excess of the lattice temperature, it becomes invalid. It has been suggested in [83] that this could be problematic in DG devices, and this is also an issue that may need addressed in further development of this methodology.

The simulator was tested using a single processor on an AMD Opteron 2.2 GHz compute cluster, and compiled using a 64-bit Intel compiler. Table 4.3 compares the average CPU time to compute a single time step for each of the quantum corrected MC simulation methodologies described above. It is clear

	Average Time Step Duration
‘Frozen Field’	0.180 secs
Classically Self Consistent	2.289 secs
Fully Self Consistent	2.334 secs

Table 4.3: Comparison of the average CPU time to carry out a time step for each method of simulation for a single test device (Device 2 in Table 5.1 in Chapter 5 with $t_{Si} = 3.3$ nm at $V_D = 0.01$ V and $V_G = 0.8$ V). The Poisson solver in the self consistent versions was Bi-CGSTAB, in each case there was an ensemble of 80 000 particles, simulated with a time step of 1×10^{-17}s. The tolerance of the solution to Poisson's equation was 1×10^{-6} where required.

that the addition of a Poisson solver adds significant compute time to the basic simulation engine (propagation, scattering and gathering of statistics) used in the ‘frozen field’ simulations. Using the same comparison of a 200 000 time step simulation, a ‘frozen field’ simulation would take around 10 hours, compared to the 3-4 days required for the self consistent equivalent. However, the addition of the Density Gradient solver does not lead to a further significant increase, making it an efficient method for self-consistent inclusion of quantum effects in MC.

4.3.5 Ohmic Contacts

The implementation of the self-consistent solution of the Density Gradient formalism in the MC simulator identified a problem concerning the injection/removal of electrons at ohmic contacts. The basic operation involves the removal of electrons that cross the boundaries at the contacts, and the injection of carriers at the contacts to match the initial fixed charge distribution obtained from the DD simulation. A k-vector is generated using a hemi-Maxwellian distribution in the direction normal to the contact, and a

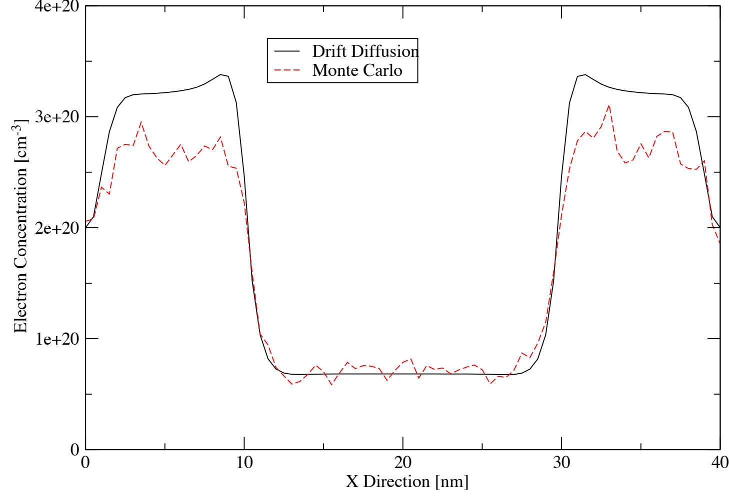

Figure 4.8: Electron concentration profile in a DG MOSFET (Device 2 in Table 5.1 in Chapter 5 with t_{Si} = 3.3 nm), running from source to drain, left to right. Shows depletion in source and drain regions that occurs when self-consistently updating the solution to the Density Gradient formalism.

Gaussian distribution in the other two directions. However, the vertical quantum distribution of particles across the channel thickness (see Fig.4.7) leads to inconsistencies. Ultimately, there was insufficient injection of carriers, especially towards the semiconductor/oxide interfaces, leading to a depletion of carriers in the source and drain regions of the device (see Fig. 4.8). This in turn leads to inconsistencies in the calculation of the quasi-Fermi level, and to a lesser extent, to the solution of Poisson's equation, thus affecting the solution of the Density Gradient equation. The self-consistent coupling of all these elements meant that the problem continually worsened as the simulation progressed.

Methods such as interpolating the carrier distribution towards the contacts or treating the contacts classically (with no quantum effects taken into account, as in [50]) were tried, but proved unsuccessful. Therefore reservoir contacts [200, 201] were implemented similar to the methodology outlined in [202].

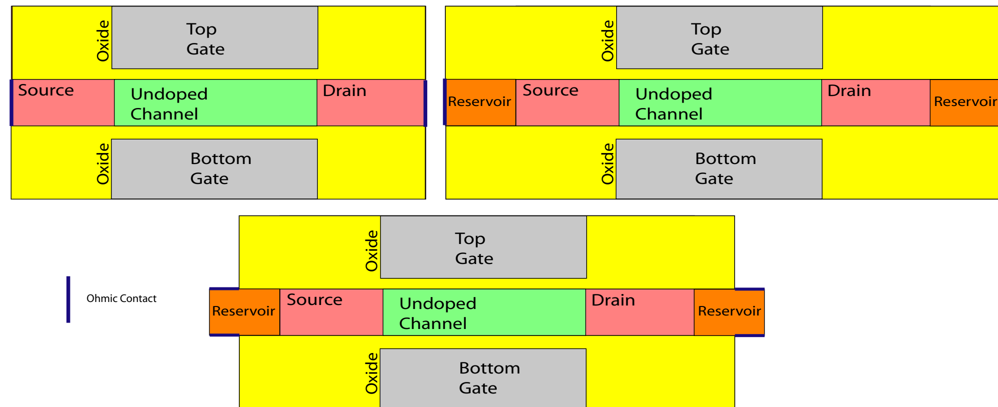

Figure 4.9: Position and orientation of ohmic contacts for a MC simulation of a DG MOSFET (arbitrary dimensions - details of the actual dimensions used in this research in Table 5.1 in Chapter 5). Clockwise from top left: the initial configuration; the first implementation of reservoir contacts; the second implementation with contacts rotated 90°.

Fig. 4.9 shows the three different placements of reservoir contacts tried in the course of this project. At the top left is the initial contact placement, covering an entire plane at the edge of the source and drain regions. This orientation was effective for classical simulations and did not cause significant problems for the quantum corrected 'frozen field' or self-consistent simulations, but as discussed above, become problematic when the quantum correction was updated in the course of the simulation.

Top right is the first implementation of reservoir contacts. The doping in these regions is an order of magnitude greater than in the source and drain regions to produce a concentration gradient and thus a diffusion current into the device. This entire structure was once again first solved in DD and, as can be seen from Fig. 4.10, there is only a negligible impact on the drain current due to

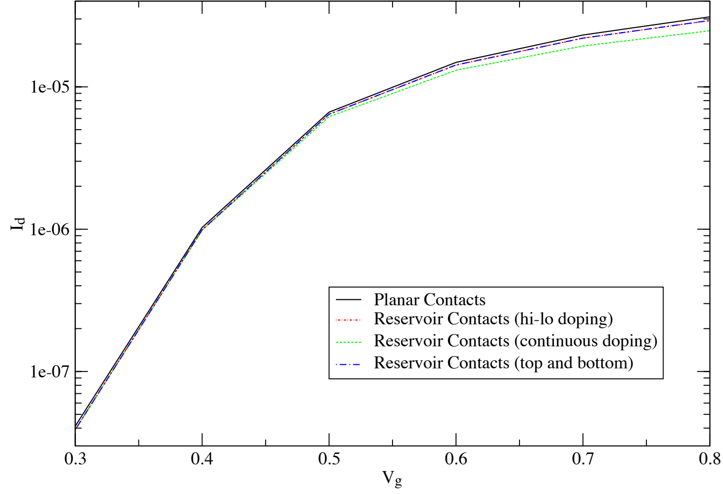

Figure 4.10: Comparison of $I_D - V_G$ characteristics for DD simulations of DG devices (dimensions without reservoirs as previously considered) with and without reservoir contacts. Three different versions of reservoir contacts were tested. The first two, with contacts aligned along the edge of the device (x-plane), first with higher doping in the reservoir regions, then again with the same doping in the reservoirs as in the source and drain. Finally, contacts orientated on the z-plane at the top and bottom of the device are shown, with higher doping in the reservoir regions.

a slight increase in series resistance. Poisson's equation was not solved in the reservoir regions, the applied field, F_{cnt}, calculated as in [202]:

$$F_{cnt} = K\Big[\int_0^{t_{Si}} N(x_{ctrl}, y)dy - \int_0^{t_{Si}} n(x_{ctrl}, y)dy\Big] \tag{4.13}$$

Here, K is a factor (dimensionless constant) used to calibrate the field to an appropriate value, and can be obtained via trial and error. Typically a value of 10^{-13} gave the most robust results. While this method showed some improvements, making the depletion less severe in the source and drain, it was still present, and ultimately led to the same problems encountered before.

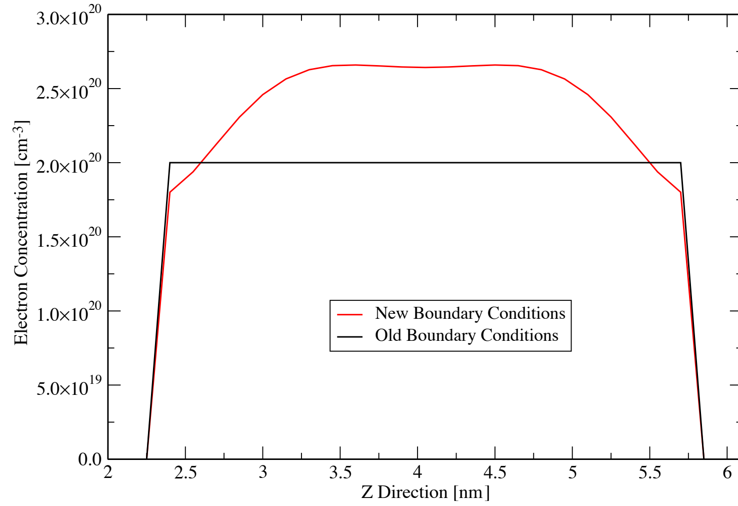

Figure 4.11: The electron concentration at the source contact in the vertical direction, as simulated using old and new boundary conditions in DD. As can be seen, with the new conditions, the typical shape associated with a quantum corrected simulation is recovered, as compared with the fixed value used in previous simulations.

The second version of the reservoir contact fix involved moving them from the x-plane, through ninety degrees onto the z-plane, on both the top and bottom of the device (see bottom diagram, Fig. 4.9). With this method, the injection was no longer influenced by the vertical quantum distribution of particles. Again, Fig. 4.10 shows this makes no significant difference to the characteristics of the device. A further difference of this method was that both Poisson and Density Gradient were solved over the entire device, including the reservoir regions. The large drawback being the increased size of the simulation domain, which now included two, significantly large reservoir regions on top of the simulated device, increasing the size of the simulated device by around 20%. Taking into account the restraints placed on the mesh spacing this led to a large grid, which, in terms of computational efficiency, is unacceptable. However, as this was being implemented and tested a more effective method was found.

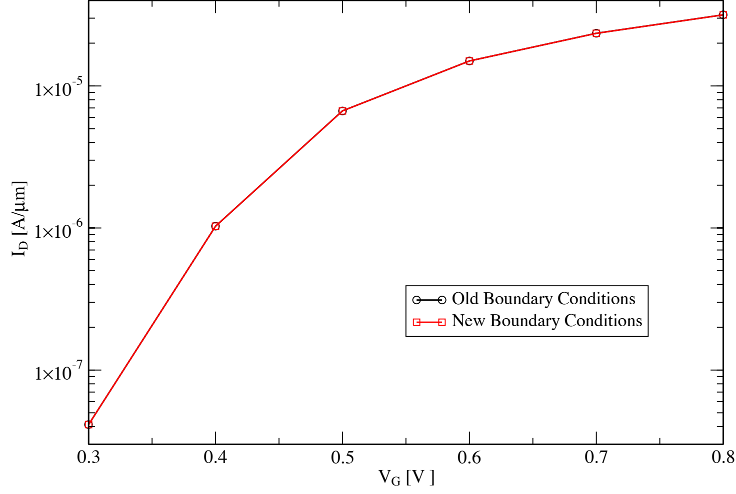

Figure 4.12: $I_D - V_G$ for a single device using both methods of boundary condition, showing there is no difference in device operation.

Ultimately, the most effective way to circumvent the contact problem was to alter the boundary conditions within the DD simulator used to initialise the MC simulator as electrons are injected into the system at the contacts consistent with the electron distribution coming from the DD solution.

Originally, Dirichlet boundary conditions were applied for the solution of Poisson, current continuity and Density Gradient equations. For Poisson's equation, the potential at the source and drain contacts and both gates was fixed consistent with the appropriate applied bias, and the carrier concentrations were fixed to a constant value for the current continuity equation. As a result, the carrier distribution used to inject electrons in the MC module was constant across the entire plane, inconsistent with the quantum corrected distribution shown in Fig. 4.7, leading to the problems described above.

To remedy this, the boundary conditions were altered to allow the potential and carrier concentration to 'float' at the source and drain contact (the conditions applied to the gates were unchanged), with the quasi-Fermi level fixed

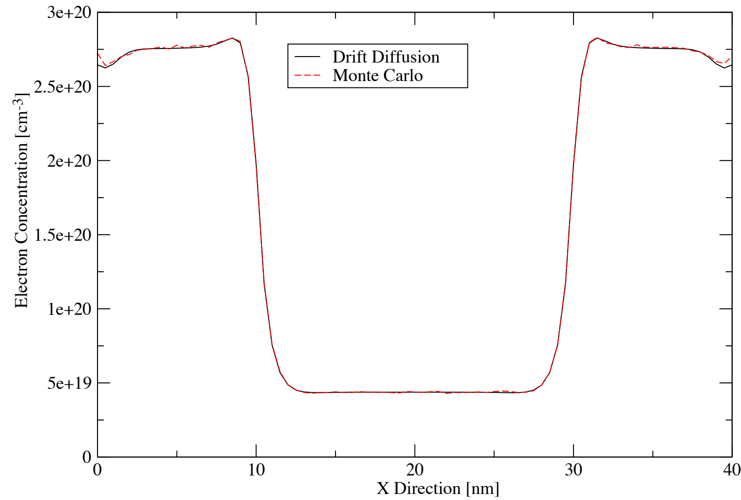

Figure 4.13: Time averaged electron concentration in the x-direction with alternate contact conditions, showing they are now well maintained, and there is no depletion present in the source and drain regions. In this instance, the time averaged electron concentration is presented in place of the more noisy, instantaneous profile.

(Neumann boundary conditions) [203]. This allowed for the electron distribution in the contacts to become consistent with the shape of the distribution in the rest of the device, and thus sufficient electrons were injected in the MC module. A comparison between the electron distribution using old and new boundary conditions at the source contact is shown in Fig 4.11, and it can be seen that the new conditions properly capture the shape associated with a quantum distribution in a UTB DG MOSFET. Fig. 4.12 demonstrates that this change in boundary conditions had no influence on the $I_D - V_G$ characteristics of the device. Fig. 4.13 shows that with the new contact conditions the problems with depletion in MC simulations are avoided and the contacts are well maintained.

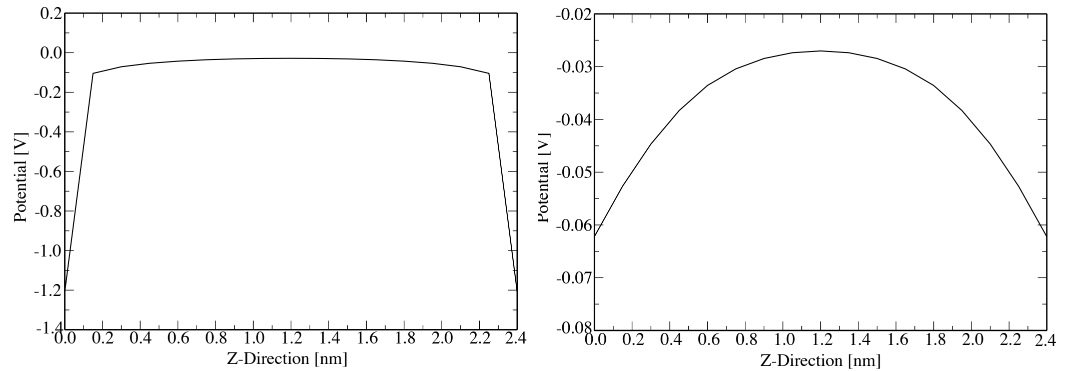

Figure 4.14: Various methods for dealing with boundary conditions associated with Density Gradient. On the left is the initial version, with concentration, and thus potential, fixed at either interface. To the right is the first fix employed in MC of interpolating the potential towards the interfaces and the gradient fix applied within the DD simulator, both of which provided an identical result. Shown here is the gradient fix.

4.3.6 Additional Considerations

The impact of the Density Gradient boundary conditions at the Si/SiO_2 interface was significant in the MC module. The initial method used in the DD code of fixing the concentration to an arbitrary low value led to discontinuity in the potential and thus a large field. This was fixed in MC by interpolating the potential towards the interfaces, thus negating the possibility of undesirable discontinuities and providing a satisfactorily smooth field. The implementation of new boundary conditions in DD, as discussed previously, meant that this was no longer necessary as the potential no longer contained undesirable discontinuities. These three methods are compared in Fig. 4.14.

As already mentioned in this chapter, there are limitations to the mesh spacings used in the MC simulator. It is desirable that the cell spacing in any direction

should not vastly exceed the Debye length (4.14) for the most heavily doped region in the simulation domain.

$$\lambda_D = \sqrt{\frac{\varepsilon k_B T}{q^2 n}} \tag{4.14}$$

It has been shown [199, 204], that using a mesh spacing of no more than $2 \times \lambda_D$ maintains stability within the simulation, and straying above that can lead to unphysical increases in carrier energy. To avoid self forces, a uniform mesh is employed, hence, in this work the mesh spacing for the x- and y-directions are restricted to 0.5 nm based on this parameter, with the spacing in the z-direction set to 0.15 nm to match the digitized steps of the interface roughness patterns, as described before.

4.3.7 Summary

In this chapter, the development of a 3D fully self-consistent MC simulator capable of capturing the effects of scattering resulting from interface roughness and body thickness variations in UTB MOSFETs was described.

The DD simulator used to provide initialization and comparison was introduced, with the most relevant sections concerning Density Gradient quantum corrections and interface roughness generation detailed more closely.

After that, the development of the 3D MC simulator was described. Initially, a 'frozen field' approximation was used, allowing for fast statistical simulations of large ensembles of devices to be carried out. The simulations, however, were limited to low V_D simulations due to lack of coupling between carrier dynamics and resulting potential distribution and electric field. The implementation of interface roughness in the MC module was detailed here as well.

To extend the validity of the simulator, a self-consistent approach was developed, coupling the electron distribution to the solution of Poisson's equation to update the classical potential coherently with the carrier propagation and the selection of an efficient BiCGSTAB solver was also justified.

As the scattering potential associated with body thickness fluctuations is related to the accompanying shift in the ground state along the channel, a method to include quantum corrections based on the Density Gradient formalism was also described. A frozen correction method was developed first, which was added to the self-consistently updated classical potential at each time step.

To extend this methodology further and allow for high V_D simulations to be carried out, a self-consistent quantum correction technique was developed via the solution of the Density Gradient equation in the MC simulator. The use of time averaging offered a compensation for the inherently noisy statistics produced by this simulation method. This periodic updating of the quantum correction term proved efficient, adding negligible simulation time overheads compared to that spent on classically self-consistent simulations, while producing accurate results.

Problems associated with the Ohmic contacts were discussed, and an eventual solution was found by altering the boundary conditions from Dirichlet to Neumann. Criteria concerning choice of mesh spacing and time step were laid out as well.

This simulator has been used to study the impact of non-uniform silicon body thicknesses in UTB SOI and DG MOSFETs. Results and discussion of these simulations are presented in the following chapter.

Chapter 5

Results And Discussion

5.1 Introduction

In this chapter results obtained using the Monte Carlo (MC) simulator developed in this project and described in Chapter 4 are presented. The main focus of these simulations was the investigation of the effect of interface roughness induced body thickness variations (BTV) on transport and variability in nanometre scale ultra thin body (UTB) MOSFETs.

As discussed in Chapter 3, previous simulation studies of BTV effects in UTB MOSFETs have been carried using the Drift-Diffusion (DD) approximation [115] which fails to account for non-equilibrium transport effects and device specific transport variations, or using Non-Equilibrium Green's Function (NEGF) [184, 185] which was restricted to 2D simulations of small devices due to the large computational effort required and does not include phonon scattering. Here the results from full 3D MC simulations (both 'frozen field' and self consistent) employing quantum corrections based on the Density Gradient formalism are presented. Firstly, the impact of BTV on mobility in long channel devices is demonstrated, showing that the 3D MC simulations can

Device	1	2	3	4
Architecture	DG	DG	DG	SOI
Top Oxide (t_{ox1}) [nm]	1.05	1.05	1.05	0.67
Bottom Oxide (t_{ox2}) [nm]	1.05	1.05	1.05	20
Silicon Thickness (t_{Si}) [nm]	varied	2.4	2.1	2.5
Source/Drain Length (L_{SD}) [nm]	10	10	10	10
Channel Length (L_{chan}) [nm]	50	20	20	10
Channel Doping [cm^{-3}]	10^{14}	10^{14}	10^{14}	10^{14}
Source/Drain Doping [cm^{-3}]	2×10^{20}	2×10^{20}	2×10^{20}	2×10^{20}

Table 5.1: The dimensions of the devices simulated in this study. These devices are illustrated in the Fig. 5.1.

capture in *ab initio* fashion the impact of the quantum confinement variation scattering on mobility. Using this as a starting point, validating the approach, the variation in on-current (I_{on}) is examined. This is an area that DD fails to accurately model, making the use of 3D MC invaluable.

5.2 Mobility Dependence on Silicon Thickness

Initially, MC simulations employing the 'frozen field' approximation (FFMC) were used to examine the impact of BTV on the mobility, μ, in DG MOSFETs. Long channeled, self-averaging devices were considered for this study, with simulations carried out at low drain voltage ($V_D = 0.05$ V), as required for 'frozen field' simulations, and for $V_G = 1.0$ V. The aim is to study the dependence of μ on silicon layer thickness (t_{Si}) in double gate (DG) MOSFETs as t_{Si} is scaled down.

The dimensions of the device indicated as Device 1 are shown in Table 5.1 and illustrated in Fig. 5.1, with t_{Si} varying from 10.8 nm down to 2.4 nm.

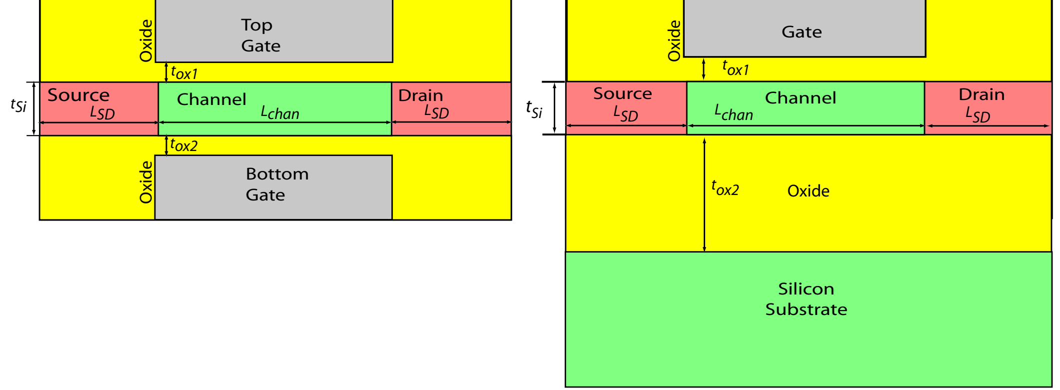

Figure 5.1: Illustration of the devices simulated in this study as detailed in Table 5.1. The DG architecture is shown on the left and SOI on the right.

The values of t_{Si} cover the range from 10 nm down to 2 nm in order to show the impact of interface roughness and BTV related scattering at the limits of scaling for these devices, where the latter scattering mechanism has been shown to have the most impact, causing significant transport degradation for $t_{Si} < 5$ nm [28]. The exact values of t_{Si} were selected in order to keep the mesh spacing constant to avoid self-force issues [196]. The mesh spacing itself was limited to multiples of the digitised interface roughness steps (0.15 nm).

For each value of t_{Si}, simulations were carried out with both uniform and non-uniform body thicknesses. The driving force is calculated using potential from an initial DD simulation employing Density Gradient quantum corrections. The inclusion of quantum corrections is important as the additional scattering due to BTV comes from the shifting of the ground state (E_0) approximated for an infinite square well by:

$$E_0 = \frac{h^2}{8m^* t_{Si}^2} \tag{5.1}$$

Where h is Plank's constant and m^* the effective mass. Density Gradient successfully approximates the shift in the ground state, as will be demonstrated

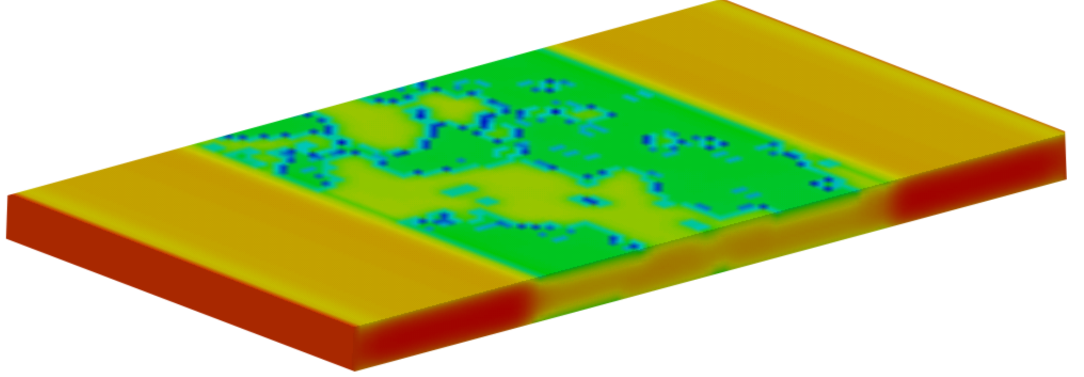

Figure 5.2: 3D plot of electron distribution in the silicon layer of Device 2 (the gates and oxide layers at the top and bottom have been removed for clarity) showing the effect of fluctuations in body thickness with the confinement along the channel. Red/orange signifies a higher concentration, green/blue signifies a lower concentration.

through the rest of this chapter. Fig. 5.2 shows the 3D electron distribution for the entire silicon layer in a thin device (t_{Si} = 2.4 nm). The top section shows the variation in electron concentration resulting from the interface roughness pattern in the channel region. Also, the fluctuation in carrier confinement along the channel can be observed, showing a clear increase in concentration in the centre of the channel when compared to the regions closer to the interfaces. The variation in confinement leads to the shifting in E_0 from equation 5.1, producing the quantum scattering potential.

Fig. 5.3 compares the potential obtained from a DD simulation with the effective quantum potential from the Density Gradient quantum corrections. As can be seen, quantum corrections are necessary to capture the fluctuations in the quantum potential that act as an additional scattering force to the carriers which is neglected in the classical version.

Figs 5.4 and 5.5 show the effect of thinning the silicon thickness on the extent of the fluctuations in the quantum potential due to BTV. Fig. 5.4 shows the

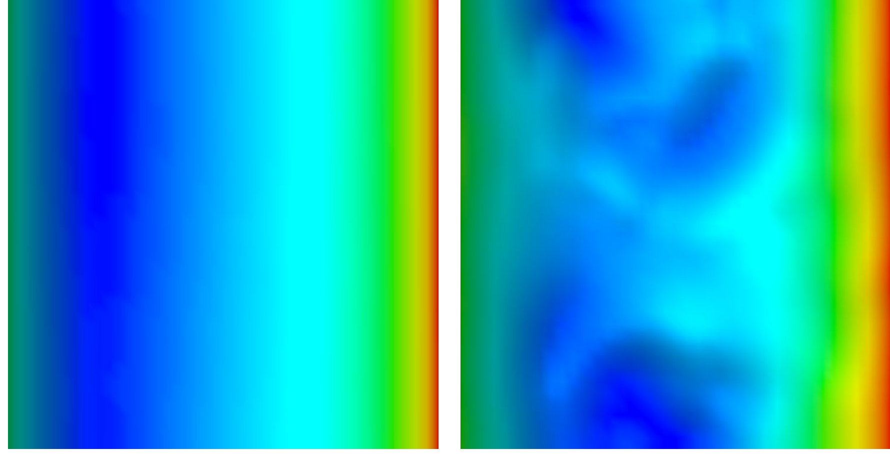

Figure 5.3: Comparison of classical (left) and quantum (right) potential plane through the centre of a 10 × 10 nm channel DG MOSFET, running from the source on the left to the drain on the right. The fluctuations due to body thickness variation are absent in the classical case, but clearly visible in the quantum corrected version. Again, red/orange signifies a higher value, green/blue signifies a lower value.

significant increase in the magnitude of the fluctuations as t_{Si} decreases, becoming most severe at the smallest t_{Si} considered. This is further emphasised by Fig. 5.5, which shows the quantum potential landscape in planes through the centre of the channel, from the source at the left to the drain at the right, for the thickest and thinnest of the devices considered here. There is a drastic difference between the smooth potential of the thickest device with the considerably rougher landscape in the thinnest device. As a result, carriers in the channel of the $t_{Si} = 2.4$ nm device experience increased scattering and thus a degradation in μ.

The μ obtained from the FFMC simulations of devices with smooth and rough interfaces for different t_{Si} are compared in Fig. 5.6, where three distinct regions can be observed.

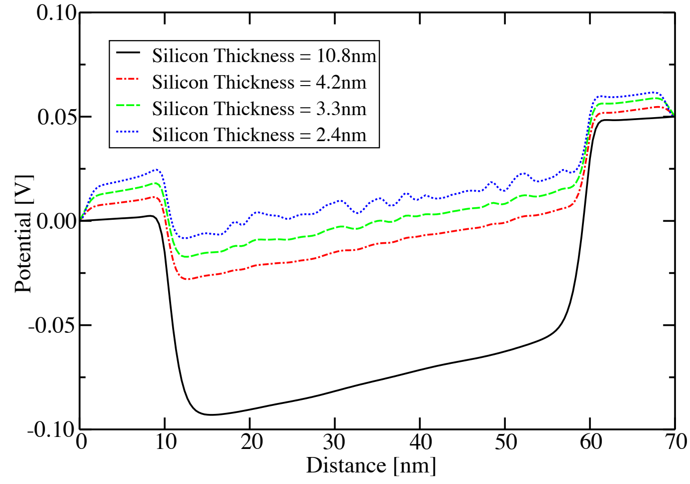

Figure 5.4: Quantum potential profile through the centre of the device, from the source at the right, to the drain at the left. Various values of t_{Si} are shown, showing greater fluctuations from variations in body thickness as the silicon layer is thinned.

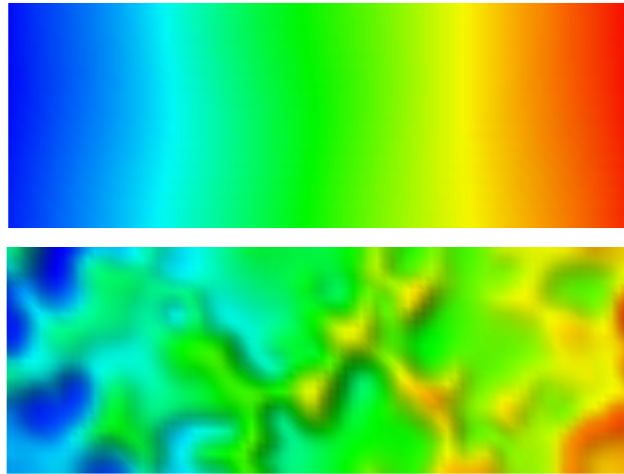

Figure 5.5: Quantum potential plane running from the source at the left to the drain at the right, taken halfway between the two interfaces. Blue represents a lower value of potential, red higher. Top is for a device with $t_{Si} = 10.8$ nm, bottom is for $t_{Si} = 2.4$ nm.

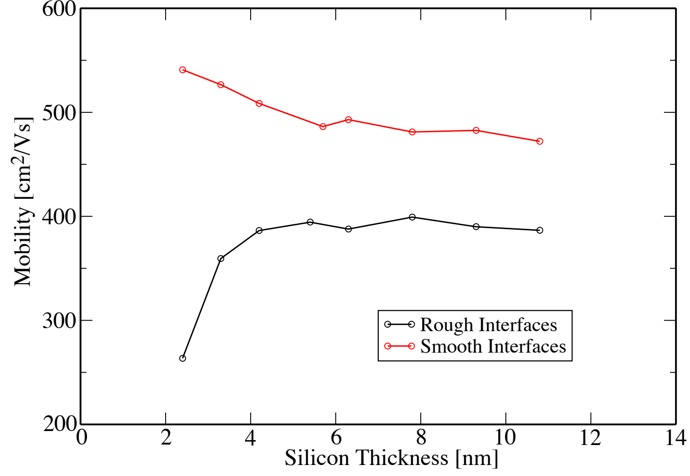

Figure 5.6: Mobility dependence on silicon thickness in DG MOSFET (Device 1 in Table 5.1).

For thick silicon layers there are two distinct channels with peak electron concentrations close to the two interfaces. The difference in mobility between the smooth and rough devices results from interface roughness scattering at the top and bottom interfaces.

As t_{Si} is reduced below 5 nm, an increase in the effective mobility for the device with uniform body thickness is observed due to volume inversion.

The sharp decrease in the mobility observed for devices with rough interfaces when $t_{Si} < 5$ nm is attributed to the increased scattering resulting from BTV as described in Chapter 2. At such channel thicknesses, the peak electron concentration moves to the centre of the channel (see Fig. 5.7), and the quantum potential fluctuations become significant through the entire silicon volume, as shown in Fig. 5.4 and 5.5, leading to mobility degradation.

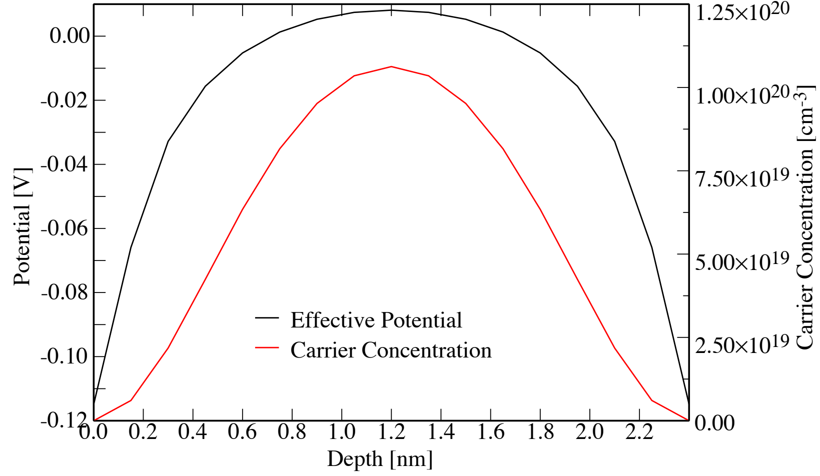

Figure 5.7: The electron distribution and quantum corrected potential for DG MOSFET with $t_{Si} = 2.4$ nm, showing in both cases the peak being in the centre of the device.

The impact of individual scattering mechanisms can be examined using Matthiessen's rule:

$$\frac{1}{\mu_{total}} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3} + ... \tag{5.2}$$

where $\mu_1, \mu_2, \mu_3, ...$ are the contributions to the total mobility (μ_{total}) of individual scattering mechanisms. In this case, taking consideration of only two mechanisms, surface roughness (μ_1) and BTV (μ_2), the effect of scattering from BTV alone on mobility can be extracted.

As shown elsewhere [78], the relationship between body thickness fluctuations limited mobility and silicon thickness is:

$$\mu \propto t_{Si}^6 \tag{5.3}$$

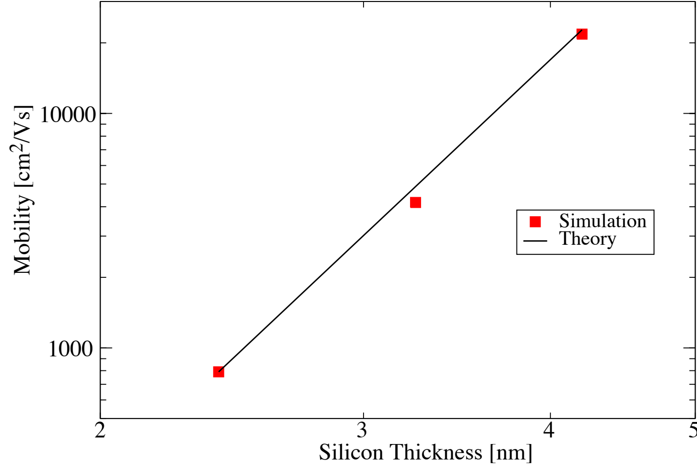

Figure 5.8: Comparison of simulated data (Fig. 5.6) to theoretical dependence of body thickness fluctuation limited mobility on t_{Si} [78].

A comparison of the mobility limited due to BTV scattering from the simulations in this work (extracted using Mathiessen's rule) in Fig. 5.6 with this theory is shown in Fig. 5.8. The very good agreement indicates that the described simulation methodology reproduces the effect of this scattering mechanism consistently with previous theoretical result [28].

The DD simulator contains a mobility model, which includes doping concentration and field dependencies. Typically, low field mobility is set using the doping concentration dependent model, the perpendicular field dependence accounts for the effects of surface roughness scattering and the lateral field dependence for velocity saturation effects. Here, for simulations of SOI and DG MOSFETs, the undoped channel, lack of interface roughness and low drain bias makes the use of this mobility model inappropriate or unnecessary for comparison of DD to MC results. So, the mobility in the DD simulator was set to a constant value, calibrated to match the mobility obtained from equivalent MC simulations (for the same device dimensions and applied bias), and this was used for the subsequent simulations to determine I_D.

5.3 Current Variation in DG and SOI MOSFETs

Having validated the methodology, showing that scattering from BTV is captured effectively using the *ab initio* technique developed in this work, the impact on the drive current (I_{on}) in UTB DG and SOI devices can be evaluated.

Here, the use of the MC rather than DD simulations is necessary for two reasons. Firstly, the strength of DD simulations is in the subthreshold region, where the current is exponentially dependent on the potential barrier between the source and the drain and the mobility plays a secondary role. It generally fails to accurately represent I_{on} due to the inability to capture non-equilibrium transport effects and transport variations due to interface roughness variations from device to device. MC exhibits the opposite behaviour. It captures well I_{on}, and its variation, but due to its statistical nature it becomes unreliable below threshold where few carriers are injected into the channel and very long simulation times are needed to gather enough statistics.

The second reason is the need to capture the scattering from BTV. DD captures only the electrostatic and quantum confinement effects which dominate the V_T variations, but it fails to capture non-equilibrium transport effects and the transport variation from device to device due to the distinct interface roughness patterns. MC is therefore a more appropriate approach to capture the BTV scattering and resulting variability as will be demonstrated through the course of this section.

5.3.1 'Frozen Field' Monte Carlo Simulation Studies

In Fig. 5.9, the $I_D - V_G$ characteristics obtained from both classical and quantum DD and FFMC simulations for Device 2 (see Table 5.1), in the absence

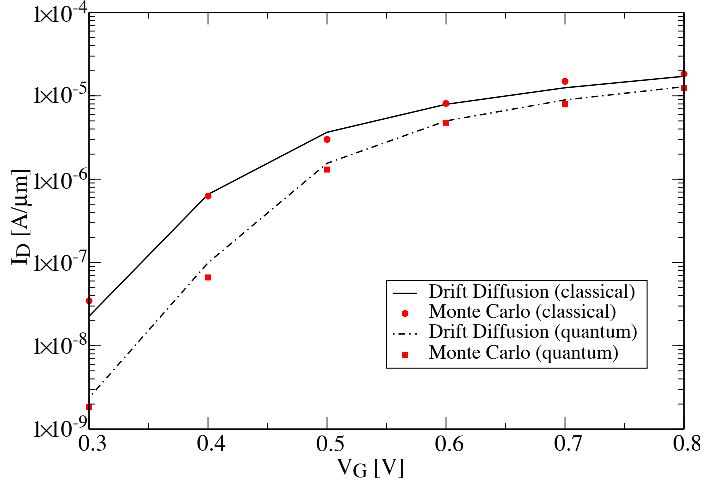

Figure 5.9: $I_D - V_G$ characteristics for DD and FFMC simulations with uniform body thickness (Device 1 in Table 5.1 with t_{Si} = 2.4 nm), comparing classical and quantum simulations, showing the associated shift in V_T.

of surface roughness are compared. Performing the simulations at low drain voltage (V_D = 1 mV) limits the non-equilibrium effects in the MC simulations and allows for fair comparison with DD simulations. As can be seen, a good agreement between the two is observed and the threshold voltage shift associated with the quantum confinement is well captured by both simulators.

An ensemble of 200 devices (Device 2 in Table 5.1) with uniquely different interface roughness patterns was simulated with the DD simulator using Density Gradient quantum corrections. Three devices were then selected, one with a high V_T, one with a low V_T, and one pitched in the middle of the distribution. Fig. 5.10 shows the resulting $I_D - V_G$ characteristics for the three devices simulated using both DD and FFMC with both uniform and non-uniform silicon body thicknesses. The DD simulations only capture the electrostatic and the quantum confinement effects, leading to a small decrease in current due to the inclusion of BTV. This comes from the shift in V_T due to the variation in thickness of the top and bottom oxide layers, and the shift in ground state (5.1) as captured by the Density Gradient quantum corrections. MC consistently

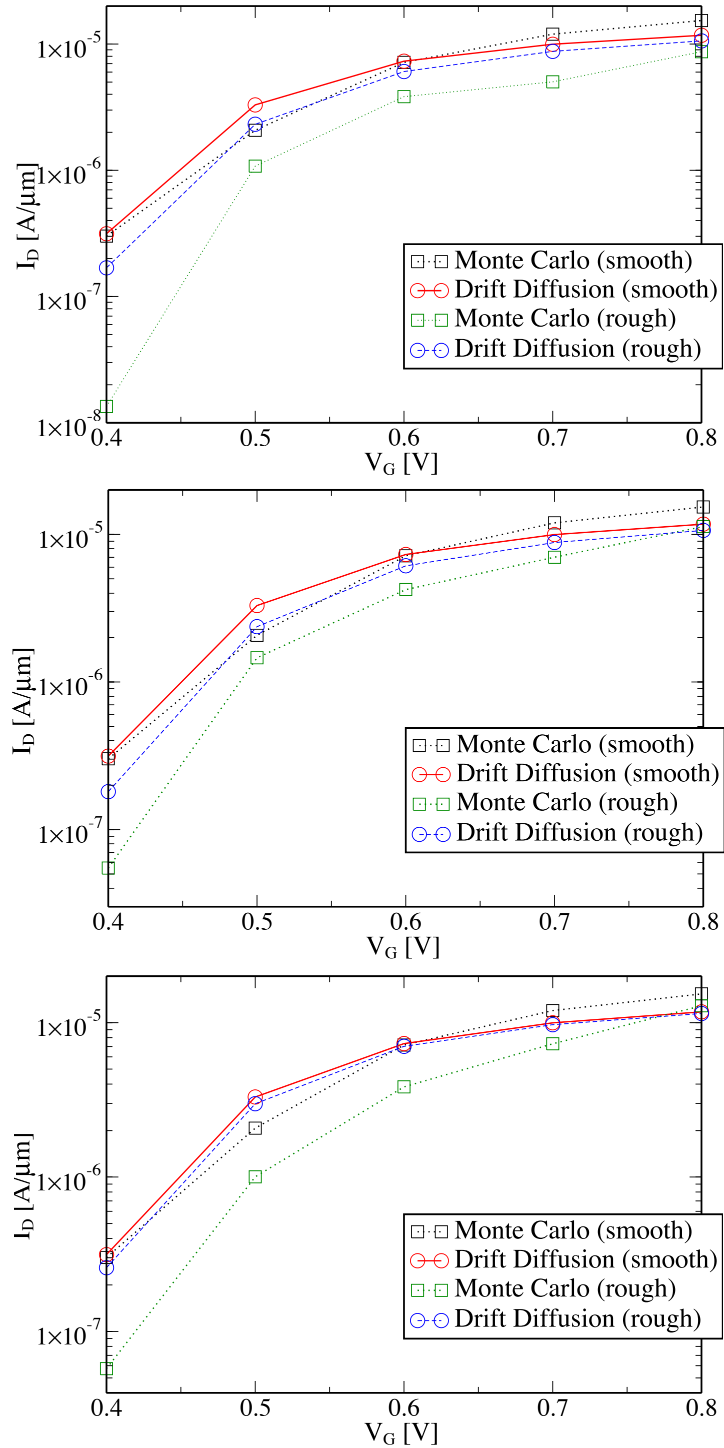

Figure 5.10: $I_D - V_G$ characteristics for DG MOSFET (Device 2 in Table 5.1) comparing DD and FFMC. $V_D = 1$ mV. Top to bottom, high, average and low V_T.

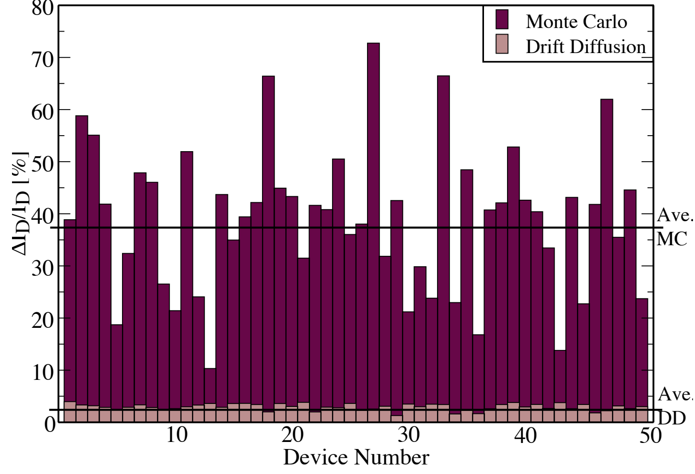

Figure 5.11: Comparison of I_D variation for 50 uniquely different DG MOSFETs, showing a greater fluctuation for MC simulations compared to DD.

captures a greater reduction in drive current due to additional scattering from body thickness fluctuations. There also appears to be no obvious relationship between V_T and the amount of degradation due to the scattering from BTV.

The next step in this study is to highlight the amount of variability that BTV introduces into UTB devices. Statistical studies using DD of large ensembles of intrinsically different devices have shown a significant V_T shift for a SOI device with t_{Si} = 5 nm due to the impact of BTV alone (without considering other sources of fluctuation) [115]. In order to attain an indication of the impact of BTV on I_{on}, similar simulation studies have been employed using MC.

50 DG devices (Device 2 in Table 5.1) with randomly generated interface roughness patterns were simulated with both DD and FFMC simulations using V_D = 1 mV, V_G = 0.8 V. A comparison of the percentage variation in I_D for each device with DD and FFMC is shown in Fig. 5.11. An identical study for a properly scaled SOI device (Device 4 in Table 5.1) was carried out, and the percentage variation in I_D is shown in Fig. 5.12. In both cases the variation

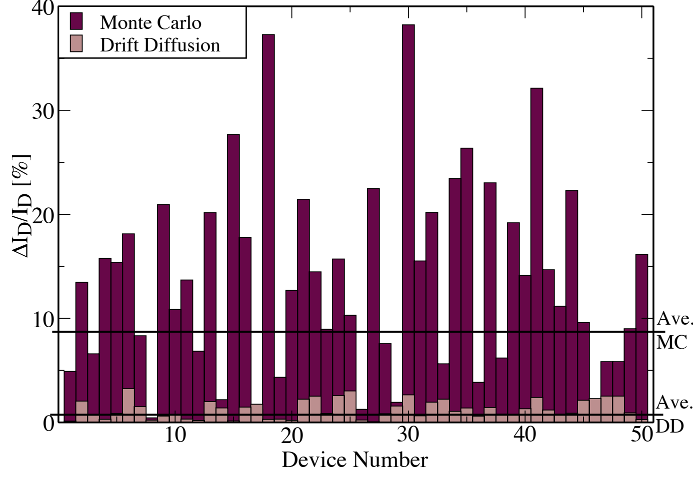

Figure 5.12: Comparison of I_D variation for 50 uniquely different SOI MOSFETs, showing a greater fluctuation for MC simulations compared to DD.

	Mean I_{on} [A/μm]	Standard Deviation [%]	Average Variation [%]
DD	1.139×10^{-5}	0.65	2.93
MC	9.387×10^{-6}	22.64	38.87

Table 5.2: Statistics from the simulation of 50 DG devices, comparing MOSFET with rough and smooth interface roughness patterns, using both DD and FFMC.

is significantly greater in MC than DD, again attributable to the inclusion of scattering from BTV.

Tables 5.2 and 5.3 compare the statistical results for the 50 devices for the DG and SOI architectures respectively. The average reduction in I_D in the long channel DG MOSFET of 38.87% is comparable to the 33.16% reduction in mobility as seen in Fig. 5.6 at the same t_{Si}. A smaller variation is shown in the properly scaled SOI device, which operates closer to the ballistic regime, though the average degradation of 8.98% is significant nonetheless, especially compared to the 0.89% obtained via DD.

	Mean I_{on} [A/μm]	Standard Deviation [%]	Average Variation [%]
DD	1.460×10^{-5}	1.05	0.89
MC	1.115×10^{-5}	11.35	8.98

Table 5.3: Statistics from the simulation of 50 SOI devices, comparing MOSFET with rough and smooth interface roughness patterns, using both DD and FFMC.

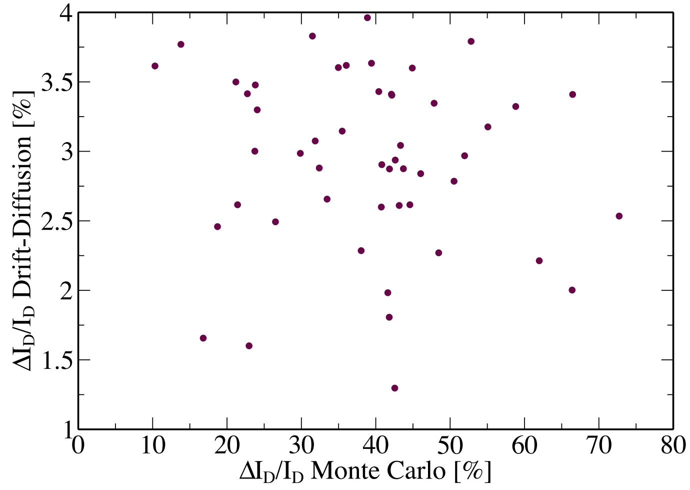

Figure 5.13: Scatter plot of I_D percentage variation for 50 uniquely different DG MOSFETs, showing weak correlation between DD and MC simulations.

Scatter plots for the percentage variation in I_D in these simulations are shown in Figs. 5.13 and 5.14 for the DG and SOI devices respectively. Extremely weak correlation is observed between the DD and FFMC simulations (the correlation coefficient is 0.105 for DG and 0.086 for SOI), which highlights the differing factors causing the current degradation and variability in each simulation methodology (electrostatics and quantum V_T shift in DD, scattering and transport variations in MC).

Therefore, scattering from BTV leads to significant variability of I_{on}. This unpredictability of the drive current from device to device resulting from addi-

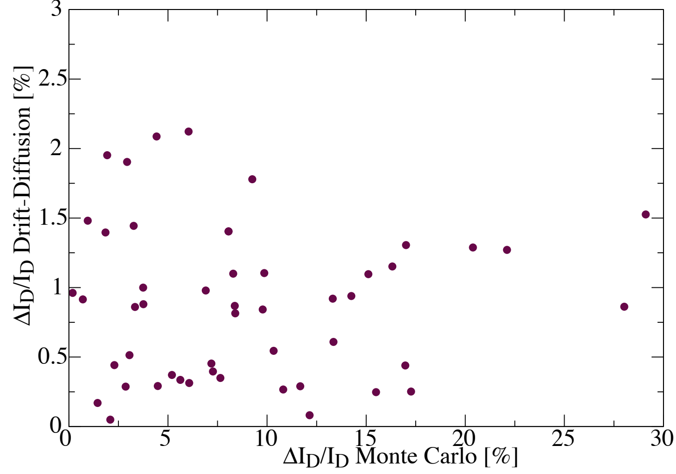

Figure 5.14: Scatter plot of I_D percentage variation for 50 uniquely different SOI MOSFETs, showing weak correlation between DD and MC simulations.

tional scattering will hamper the integration into future chips where they will number in the billions.

5.3.2 Self-Consistent Monte Carlo Simulation Studies

The inclusion of a Poisson solver, as detailed in Chapter 4, allows for the electric field to be updated to reflect changes in the mobile carrier distribution. This method means classical simulations can be carried out at higher drain bias, however, the method for inclusion of frozen quantum corrections again restricts quantum simulations to a lower drain bias. Fig. 5.15 shows the smooth classical potential obtained from the Poisson solver employed in the MC module. The MC simulator can now be run in a self-consistent mode with a frozen quantum correction (FQMC).

There is good agreement between the three quantum corrected simulation methods (DD with Density Gradient, FFMC and FQMC), as shown in Fig. 5.16

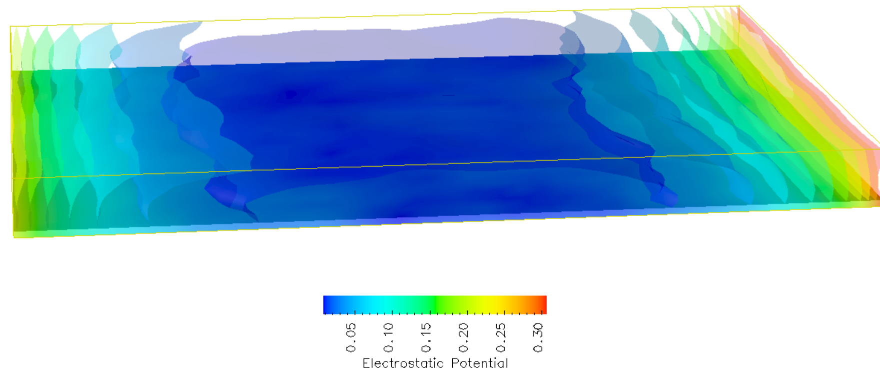

Figure 5.15: 3D plot of classical potential, bottom half of the channel, running left to right from the source to the drain, at $V_D = 1$ mV, $V_G = 0.8$ V, with the units also in Volts.

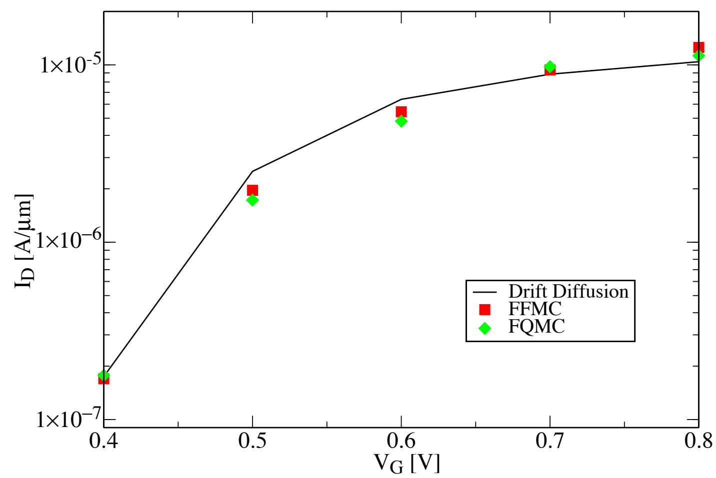

Figure 5.16: $I_D - V_G$ characteristics for DG MOSFET (Device 3 in Table 5.1) in the absence of interface roughness comparing DD, FFMC and FQMC, all including Density Gradient quantum corrections and at $V_D = 1$ mV .

	Drift-Diffusion	Monte Carlo
I_D smooth interfaces [A/μm]	1.042×10^{-5}	1.195×10^{-5}
I_D rough interfaces [A/μm]	9.59×10^{-6}	8.74×10^{-6}
Variation [%]	7.92	26.87

Table 5.4: Comparison of quantum corrected DD and FQMC simulations of Device 3 at $V_D = 1$ mV, $V_G = 0.8$ V.

which compares $I_D - V_G$ characteristics in the absence of interface roughness patterns.

Fig. 5.17 shows the electron concentration in the channel of Device 2 in Table 5.1, taken in a vertical plane, from interface to interface. This is from a FQMC simulation without interface roughness patterns and shows the peak carrier concentration to again be confined to the centre of the channel.

The introduction of interface roughness produces the variation in confinement demonstrated in Fig. 5.18 for a DD simulation and Fig. 5.19 for FQMC simulation of the same device (Device 3 in Table 5.1). In both cases similar confinement patterns are observed, though the DD plot is smoother due to noise present in the MC statistics. Comparing the classical potential in Fig. 5.15 where the potential is smooth with the quantum potential in Fig. 5.20 where fluctuations are prominent, shows how effectively the quantum corrections in this methodology capture the additional scattering potential.

This results in fluctuation in the electron distribution as shown in Fig. 5.21, leading to degradation in transport. A comparison between simulations carried out using this method and DD is shown in Table 5.4, again demonstrating that a greater variation in I_D is captured with FQMC simulations due to the additional scattering. Also the variation captured (26.87%) is of a similar magnitude to the reduction seen in both mobility (33.16%) and I_{on} (38.87%)

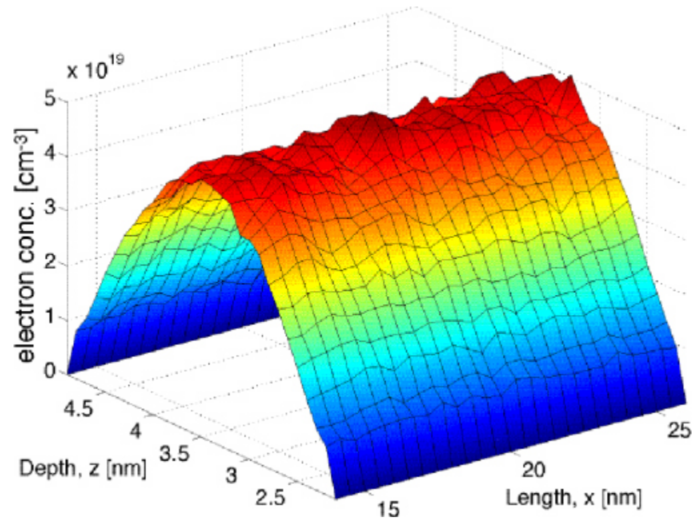

Figure 5.17: MC carrier concentration on a vertical plane in the channel of a DG MOSFET in the absence of interface roughness patterns.

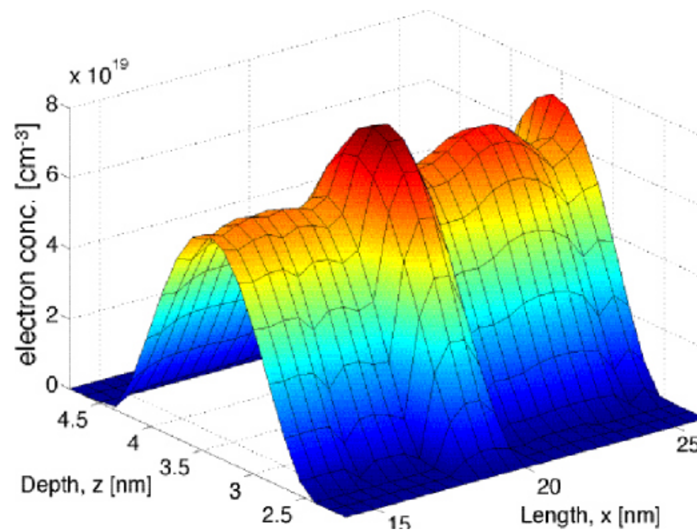

Figure 5.18: DD carrier concentration on a vertical plane in the channel of a DG MOSFET in the presence of roughness patterns at both interfaces. Variations in concentration due to confinement can be observed.

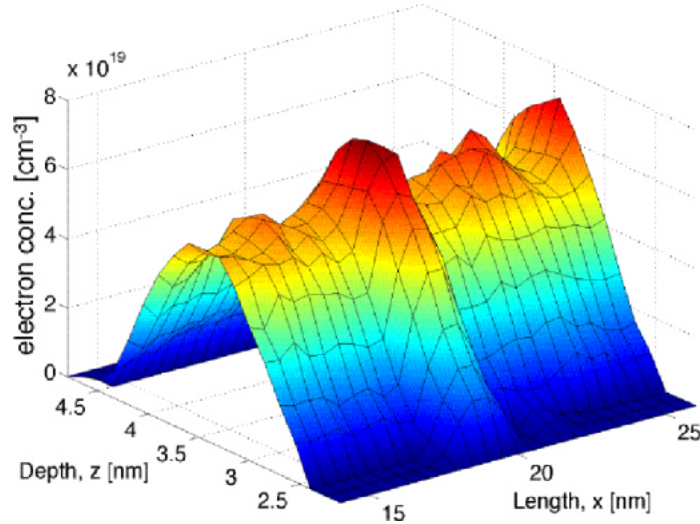

Figure 5.19: MC carrier concentration on a vertical plane in the channel of a DG MOSFET in the presence of roughness patterns at both interfaces. Variations in concentration due to confinement can be observed.

in the FFMC simulations. A study of a larger number of devices may improve the comparison further.

In these low field simulations, the variation of I_D is greater than would be expected at high field as the transport is mobility dominated. Hence, the use of an ensemble MC simulator is necessary, as these scattering effects are not accounted for in the DD approximation.

In order to carry out quantum corrected simulations at high V_D, it was necessary to introduce a Density Gradient solver to make the simulator fully self-consistent (SCQMC - the development of this approach was detailed in Chapter 4). At higher V_D, the importance of quantum confinement scattering decreases as the degree of balisticity increases so less variation in I_{on} would be expected. The problems associated with the employment of this methodology were described in Chapter 4, along with the solutions leading to the eventual successful implementation. The time spent developing the code left little

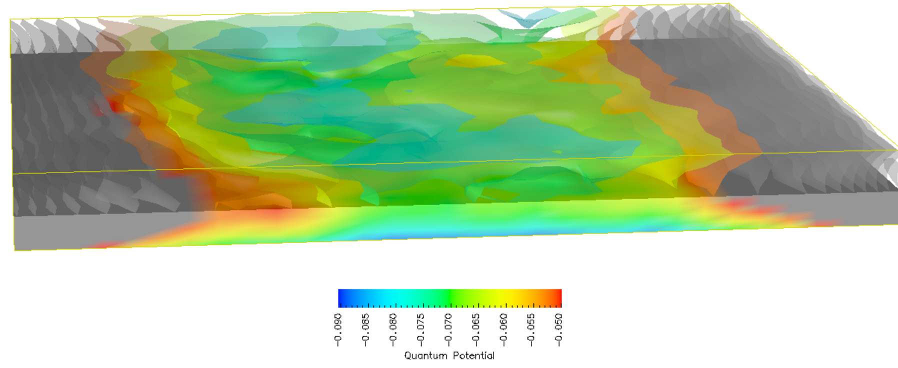

Figure 5.20: 3D plot of quantum corrected potential, as calculated in the MC simulator, at V_D = 1 mV, V_G = 0.8V, with the units also in Volts.

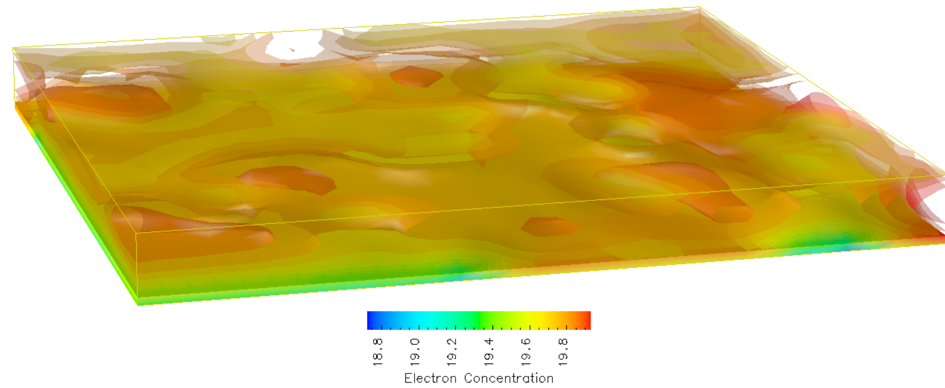

Figure 5.21: 3D plot of electron concentration in Device 2 (Table 5.1) from a MC simulation, running from the source to the drain, left to right. The units here are $\times 10^{19}\text{cm}^{-3}$.

time to run substantial simulations with the fully self consistent MC simulator at high V_D, so further validation of this methodology, along with high field simulations are still to be carried out.

5.4 Summary

The variability of I_{on} due to BTV in UTB SOI and DG MOSFETs has been investigated using a 3D MC simulator employing quantum corrections based on the Density Gradient formalism. Scattering from surface roughness and BTV is included in the *ab initio* approach described in Chapter 4.

Initially, a 'frozen field' approximation was employed to study mobility dependence on silicon layer thickness in long channeled DG MOSFETs. In this instance, the field was generated using an initial DD simulation, and never updated through the course of the MC simulation. For 5 nm $< t_{Si} <$ 10 nm, the observed degradation in μ is attributed to surface roughness scattering. Below 5 nm, an increase in μ for devices with uniform silicon thicknesses is observed as a result of volume inversion. For non-uniform silicon thicknesses, a drop in μ is captured as a result of scattering from BTV. This is shown to be consistent with the t_{Si}^6 dependence of μ predicted by theory [28], which validates the use of this approach for studying BTV scattering.

The next step was to employ the MC simulator to study the variation of I_{on} in UTB devices. The failure of DD simulations to properly capture current above threshold as well as the impact of scattering from BTV makes the use of MC simulations vital. Large ensembles of devices were simulated to highlight two aspects of this phenomena. Firstly, the average variation in I_D is considerably greater in MC than DD, as DD is incapable of capturing the scattering associated with BTV. Additionally, by comparing the variation in sets of devices with uniquely different roughness patterns, the vastly differing degrees of I_{on} degradation in each individual device has been shown. This device to device

variability will hamper the integration of devices using this architecture into future circuits where they will number in the billions.

The development of self consistency of the MC simulator with Poisson's equation and with Density Gradient will allow for a greater range of simulations to be carried out, though, at this time the full range of possibilities using this methodology have yet to be explored.

Chapter 6

Conclusions

The aim of this thesis was the development of a simulation methodology and tools that accurately and efficiently capture the impact of interface roughness related silicon body thickness variations in ultra thin body (UTB) silicon on insulator (SOI) double gate (DG) MOSFETs at the limits of device scaling. A 3D Monte Carlo (MC) simulator was developed that was used, along side a 3D Drift Diffusion (DD) simulator, to examine the degradation in transport due to this scattering mechanism. The significance of this phenomena was highlighted by use of statistical studies showing the current variability from device to device due to the uniqueness of the body thickness variation in each transistor. This will adversely effect the integration of ultimate UTB SOI and DG MOSFETs into modern and future chips where they number in the billions.

Chapter 2 discussed the advantages offered in scaling SOI and DG MOSFETs in comparison to their conventional counterparts. The geometry of these device counteracting short channel effects, and the virtually undoped channel allowing for higher carrier mobility. The complex transport behaviour associated with these architectures was discussed, detailing research carried out previously. The transport enhancement associated with volume inversion and band splitting was contrasted with the detrimental effects of scattering from

Coulomb interactions with charges trapped at the interfaces, phonons (due to confinement of carriers and acoustic phonons), interface roughness and body thickness variations. The last of these was of particular interest in this study, becoming extremely significant when $t_{Si} < 5$ nm, with its effects on transport being the main focus of this work.

In Chapter 3, the selection of an appropriate simulation methodology to examine the quantum confinement scattering and its impact on device variability was discussed. MC, DD and Non-Equlibrium Green's Functions (NEGF) were all described and evaluated in terms of their suitability. The computational efficiency of DD is attractive, allowing for large statistical studies to be carried out quickly, and is capable of including quantum effects, but its effectiveness is limited by its inability to capture non-equilibrium effects and the transport variations from device to device due to the scattering from the unique body thickness fluctuations which is of significant interest in this work. The NEGF formalism offers the reverse scenario - while it can capture the transport variability associated with this scattering mechanism, the computational efficiency rapidly decreases as the complexity of the model (such as moving from two to three dimensions or the inclusion of complex scattering mechanisms) increases, generally limiting 3D simulations to small structures which was unsuitable for this work where large self averaging devices were under consideration. Hence, 3D MC was deemed to be the most appropriate simulation methodology for this research. It remains computationally efficient, even when extended to 3D, is capable of capturing non-equilibrium carrier transport and can include quantum effects through the introduction of a quantum corrected potential with a minimal computational cost. Various methods of implementing quantum corrections in MC were compared, based on the Schrödinger equation, the Wigner distribution function, the Böhm interpretation of the Schrödinger equation (Density Gradient and Effective Conduction Band Edge) and effective potential approaches. The derivations of these formalisms and the relative strengths and weaknesses were also discussed.

Chapter 4 describes the development of the 3D MC simulator used for this study. First, the DD simulator used to initialize and validate the MC module was described, including details of the surface roughness generation technique used to create the semiconductor/oxide interface roughness patterns and resulting non-uniform body thickness. Following that, a detailed description of the MC module was presented. General details of the MC model were described, in terms of charge assignment scheme, band structure model and existing scattering mechanisms. The 'frozen field' approximation was introduced, followed by the description of the self-consistent version of the simulator. The quantum correction scheme, based on the Density Gradient formalism was then introduced. The importance of the inclusion of quantum effects comes as the body thickness variation induced scattering is a product of a random quantum confinement variations. Therefore, the quantum corrections allow for an *ab initio* treatment of interface roughness and body thickness fluctuation induced scattering to be employed. This chapter also dealt with issues associated with the boundary conditions at the source and drain contacts, which provided a major problem in the development of the MC module. The existing Dirichlet boundary conditions lead to insufficient injection of carriers into the device when quantum corrections were introduced, and therefore depletion in the source and drain regions. Self-consistent coupling of the current to Poisson's equation and the Density Gradient equation exacerbated this problem further. The implementation of Neumann boundary conditions resulted in the contacts being properly maintained, remedying this issue. As a result, the first 3D MC simulator capable of studying quantum confinement transport variations has been successfully developed.

Finally, in Chapter 5, results from statistical simulation studies carried out using this simulation methodology were presented, using both the 'frozen field' and classically self-consistent with a frozen quantum correction versions of the MC module. The simulator was shown to capture well the quantum confinement related t_{Si}^6 mobility dependence on silicon thickness predicted by theory [28], validating the *ab initio* inclusion of the additional scattering from

body thickness variations. The impact on the drive current in these devices was demonstrated via careful comparison with DD. Good agreement between the $I_D - V_G$ characteristics from DD and MC simulations at low drain bias (where non-equilibrium effects are limited), and in the absence of interface roughness patterns, was demonstrated. When interface roughness patterns were introduced, the greater degradation in I_D exhibited by the MC simulations is attributed to the additional scattering captured by MC, but neglected by DD where only electrostatic effects are accounted for. Simulations of large ensembles of devices with uniquely different roughness patterns showed for the first time a large on-current variation from device to device due to the additional quantum confinement scattering. This level of unpredictability will make the integration of ultimately scaled UTB SOI and DG MOSFETs into modern and future circuits problematic.

6.1 Suggestions for Future Work

Further studies, most specifically at higher drain bias, could be carried out using the existing simulator to provide further validation of the methodology. Simulations of realistic modern or future devices using the MC simulator would also be of considerable interest. Less variation in I_D would be expected here, as transport becomes less influenced by scattering and more ballistic.

In order to gain a more comprehensive and accurate picture of transport phenomena in UTB SOI and DG MOSFETs described in Chapter 2, the MC simulator could be extended yet further. The first extension, would be to move from the spherical band structure to an ellipsoidal one [120], allowing for the capture of the band splitting phenomena as the silicon thickness is scaled [67].

The additional scattering mechanisms detailed in Chapter 2 could be implemented by employing the approaches used by others. Coulomb scattering could

be captured using an *ab initio* approach as described in Chapter 3, consistent with the methodology detailed in [143] and the increased phonon scattering accounted as in [53] and [57]. From this a full, 3D MC simulator capable of accurately simulating the behavior and current variability in UTB architectures down to the limits of scaling would have been developed.

Another possible extension would be to introduce a non-equilbrium carrier temperature distribution in the Density Gradient solver. As it stands, the thermal equilibrium approximation is employed, assuming the carriers are kept at the same temperature as the lattice (typically room temperature, $T = 300$ K). A methodology has been proposed in [160] for the Effective Conduction Band Edge (ECBE) formalism, where the $k_B T$ term is replaced by the average value of:

$$U = \langle v_z \hbar k_z \rangle \tag{6.1}$$

Where v_z is the velocity and k_z the momentum in the z-direction. Taken in slices along the x-direction, this would account for changes in the carrier temperature. The authors in [160] have shown this to be an effective method to deal with this issue, and it could perhaps be included within the Density Gradient solver used in this study in a similar manner. This would allow for accurate simulation where high fields result in the temperature of the carriers exceeding that of the lattice.

Bibliography

[1] T. Skotnicki, "Materials and device structures for sub-32 nm CMOS nodes," *Microelectronic Engineering*, vol. 84, pp. 1845–1852, 2007.

[2] T. Vogelsang and K. R. Hofmann, "Electron transport in strained Si layers on $\mathrm{Si_{1-x}Ge_x}$ substrates," *Applied Physics Letters*, vol. 63, no. 2, pp. 186–188, July 1993.

[3] K. Ismail, J. O. Chu, and B. S. Meyerson, "High hole mobility in SiGe alloys for device applications," *Applied Physics Letters*, vol. 64, no. 23, pp. 3124–3126, June 1994.

[4] M. V. Fischetti and S. E. Laux, "Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys," *Journal of Applied Physics*, vol. 80, no. 4, pp. 2234–2252, August 1996.

[5] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Transconductance Enhancement in Deep Submicron Strained-Si n-MOSFETs," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 707–710, 1998.

[6] J. R. Watling, L. Yang, M. Boriçi, R. C. W. Wilkins, A. Asenov, J. R. Barker, and S. Roy, "The Impact of interface roughness scattering and degeneracy in relaxed and strained Si n-channel MOSFETs," *Solid-State Electronics*, vol. 48, pp. 1337–1346, 2004.

[7] H.-S. P. Wong, "Beyond the conventional transistor," *IBM Journal of Research And Development*, vol. 46, no. 2/3, pp. 133–168, March/May 2002.

[8] J. L. Hoyt, H. M. Nayfeh, S. Eguchi, I. Aberg, G. Xia, T. Drake, E. A. Fitzgerald, and D. A. Antoniadis, "Strained Silicon MOSFET Technology," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 23–26, 2002.

[9] M. Ieong, B. Doris, J. Kedzierski, K. Rim, and M. Yang, "Silicon Device Scaling to the Sub-10-nm Regime," *Science*, vol. 306, pp. 2057–2060, December 2004.

[10] K. Uchida, R. Zednik, C.-H. Lu, H. Jagannathan, J. McVittie, P. C. McIntyre, and Y. Nishi, "Experimental Study of Biaxial and Uniaxial Strain Effects on Carrier Mobility in Bulk and Ultrathin-body SOI MOSFETs," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 229–232, 2004.

[11] K. Mistry, M. Armstrong, C. Auth, S. Cea, T. Coan, T. Ghani, T. Hoffmann, A. Murthy, J. Sandford, R. Shaheed, K. Zawadzki, K. Zhang, S. Thompson, and M. Bohr, "Delaying Forever: Uniaxial Strained Silicon Transistors in a 90nm CMOS Technology," *Digest of Technical Papers, 2004 Symposium on VLSI Technology*, pp. 50–51, 2004.

[12] Y. C. Wang, M. Hong, J. M. Kuo, J. P. Mannaerts, J. Kwo, H. S. Tsai, J. J. Krajewski, Y. K. Chen, and A. Y. Cho, "Demonstration of Submicron Depletion-Mode GaAs MOSFETs with Negligible Drain Current Drift and Hysteresis," *IEEE Electron Device Letters*, vol. 20, no. 9, pp. 457–459, September 1999.

[13] K. Kalna, L. Yang, and A. Asenov, "Monte Carlo simulations of sub-100 nm InGaAs MOSFETs for digital applications," *Proceedings of the 35th European Solid-State Device Research Conference (ESSDERC)*, pp. 169–172, 2005.

[14] S. Datta, T. Ashley, J. Brask, L. Buckle, M. Doczy, M. Emeny, D. Hayes, K. Hilton, R. Jefferies, T. Martin, T. J. Phillips, D. Wallis, P. Wilding, and R. Chau, "85nm Gate Length Enhancement and Depletion mode InSb Quantum Well Transistors for Ultra High Speed and Very Low Power Digital Logic Applications," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 763–766, 2005.

[15] R. Droopad, M. Passlack, N. England, K. Rajagopalan, J. Abrokwah, and A. Kummel, "Gate dielectrics on compound semiconductors," *Microelectronic Engineering*, vol. 80, pp. 138–145, 2005.

[16] M. V. Fischetti, S. E. Laux, P. M. . Solomon, and A. Kumar, "Thirty Years of Monte Carlo Simulations of Electronic Transport in Semiconductors: Their Relevance to Science and Mainstream VLSI Technology," *Journal of Computational Electronics*, vol. 3, pp. 287–293, 2004.

[17] A. Pethe, T. Krishnamohan, D. Kim, S. Oh, H.-S. P. Wong, Y. Nishi, and K. C. Saraswat, "Investigation of the Performance limits of III-V Double-Gate n-MOSFETs," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 605–608, 2005.

[18] K. Kalna, A. Asenov, and M. Passlack, "Monte Carlo Simulation of Implant Free InGaAs MOSFET," *Journal of Physics: Conference Series*, vol. 38, pp. 200–203, 2006.

[19] M. Passlack, R. Droopad, K. Rajagopalan, J. Abrokwah, P. Zurcher, and P. Fejes, "High Mobility III-V MOSFET Technology," *IEEE Compound Semiconductor Integrated Circuit Symposium*, pp. 39–42, 2006.

[20] A. I. Kingon, J.-P. Maria, and S. K. Streiffer, "Alternative dielectrics to silicon dixode for memory and logic devices," *Nature*, vol. 406, pp. 1032–1038, August 2000.

[21] M. V. Fischetti, D. A. Neumayer, and E. A. Cartier, "Effective electron mobility in Si Inversion laters in metal-oxide-semiconductor systems with

a high-κ insulator: The role of remote phonon scattering," *Journal of Applied Physics*, vol. 90, no. 9, pp. 4587–4608, November 2001.

[22] R. Chau, S. Datta, M. Doczy, B. Doyle, J. Kavalieros, and M. Metz, "High-κ/Metal-Gate Stack and Its MOSFET Characteristics," *IEEE Electron Device Letters*, vol. 25, no. 6, pp. 408–410, June 2004.

[23] J.-P. Colinge, "Multiple-gate SOI MOSFETs," *Solid-State Electronics*, vol. 48, pp. 897–905, 2004.

[24] A. Kranti and G. A. Armstrong, "Optimization of the source/drain region profile for suppression of short channel effects in sub-50 nm DG MOSFETs with high-κ gate dielectrics," *Semiconductor Science and Technology*, vol. 21, pp. 1563–1572, 2006.

[25] D. A. Antoniadis, I. Aberg, C. N. Chléirigh, O. M. Nayfeh, A. Kakifirooz, and J. L. Hoyt, "Continuous MOSFET performance increase with device scaling: The role of strain and channel material innovations," *IBM Journal of Research And Development*, vol. 50, no. 4/5, pp. 363–376, July/September 2006.

[26] C. Fenouillet-Beranger, T. Skotnicki, S. Monfray, N. Carriere, and F. Boeuf, "Requirements for ultra-thin-film devices and new materials for the CMOS roadmap," *Solid-State Electronics*, vol. 48, pp. 961–967, 2004.

[27] ITRS, *International Technology Roadmap for Semiconductors, 2005 Edition.* Semiconductor Industry Association (SIA), 2005.

[28] K. Uchida and S. Takagi, "Carrier scattering induced by thickness fluctuation of silicon-on-insulator film in ultrathin-body metal-oxide-semiconductor field-effect transistors," *Applied Physics Letters*, vol. 82, no. 17, pp. 2916–2918, April 2003.

[29] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of Ion-Implanted MOSFET's with Very

Small Physical Dimensions," *IEEE Journal of Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, October 1974.

[30] C. Fiegna, H. Iwai, T. Wada, M. Saito, E. Sangiorgi, and B. Riccò, "Scaling the MOS Transistor Below 0.1 μm: Methodology, Device Structures, and Technology Requirements," *IEEE Transactions on Electron Devices*, vol. 41, no. 6, pp. 941–951, June 1994.

[31] Y. Taur, D. A. Buchanan, W. Chen, D. J. Frank, K. E. Ismail, S.-H. Lo, G. A. S. Halasz, R. G. Viswanathan, H.-J. C. Wann, S. J. Wind, and H.-S. Wong, "CMOS Scaling into the Nanometer Regime," *Proceedings of the IEEE*, vol. 85, no. 4, pp. 486–504, April 1997.

[32] D. L. Critchlow, "MOSFET Scaling - The Driver of VLSI Technology," *Proceedings of the IEEE*, vol. 87, no. 4, pp. 659–667, April 1999.

[33] Y. Taur, C. H. Wann, and D. J. Frank, "25 nm CMOS Design Considerations," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 789–792, 1998.

[34] B. Yu, H. Wang, A. Joshi, Q. Xiang, E. Ibok, and M. R. Lin, "15 nm Gate Length Planar CMOS Transistor," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 937–941, 2001.

[35] C.-T. Chuang, K. Bernstein, R. V. Joshi, R. Puri, K. Kim, E. J. Nowak, T. Ludwig, and I. Aller, "Scaling Planar Sillicon Devices," *IEEE Circuits And Devices Magazine*, pp. 6–19, January/February 2004.

[36] Y. Taur and T. H. Ning, *Fundamentals of Modern VSLI Devices.* Cambridge University Press, Cambridge, 1998.

[37] A. Asenov and S. Saini, "Suppression of Random Dopant-Induced Threshold Voltage Fluctuations in Sub-0.1-μm MOSFET's with Epitaxial and δ-Doped Channels," *IEEE Transactions on Electron Devices*, vol. 46, no. 8, pp. 1718–1724, August 1999.

[38] L. Chang, Y.-K. Choi, D. Ha, P. Ranade, S. Xiong, J. Bokor, C. Hu, and T.-J. King, "Extremely Scaled Silicon Nano-CMOS Devices," *Proceedings of the IEEE*, vol. 91, no. 11, pp. 1860–1873, November 2003.

[39] H. P. Wong, D. J. Frank, P. M. . Solomon, C. H. J. Wann, and J. J. Welser, "Nanoscale CMOS," *Proceedings of the IEEE*, vol. 87, no. 4, pp. 537–570, 1999.

[40] M. M. Pelella, W. Maszara, S. Sundararajan, S. Sinha, A. Wei, D. Ju, W. En, S. Krishnan, D. Chan, S. Chan, P. Yeh, M. Lee, D. Wu, M. Fuselier, R. van Bentum, G. Burdach, C. Lee, G. Hill, D. Greenlaw, C. Riccobene, O. Karlsson, D. Wristers, and N. Kepler, "Advantages and Challenges of High Perfomrance CMOS on SOI," *IEEE International SOI Conference*, pp. 1–4, 2001.

[41] J. Saint-Martin, A. Bournel, and P. Dollfus, "Comparison of multiple-gate MOSFET architectures using Monte Carlo simulation," *Solid-State Electronics*, vol. 50, pp. 94–101, 2006.

[42] C. Fiegna, H. Iwai, T. Wada, T. Saito, E. Sangiorgi, and B. Riccò, "A New Scaling Methodology for the 0.1 - 0.025 μm MOSFET," *Symposium on VLSI Technology Digest of Technical Papers*, pp. 33–34, 1993.

[43] H.-S. P. Wong, D. J. Frank, and P. M. Solomon, "Device Design Consideration for Double-Gate, Ground-Plane, and Single-Gated Ultra-Thin SOI MOSFET's at the 25nm Channel Length Generation," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 407–410, 1998.

[44] D. Frank, S. Laux, and M. Fischetti, "Monte Carlo Simulations of a 30 nm Dual-Gate MOSFET: How Short Can Si Go?" *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 553–556, 1992.

[45] W. Chaisantikulwat, M. Mouis, G. Ghibaudo, S. Cristoloveanu, J. Widiez, M. Vinet, and S. Deleonibus, "Experimental Evidence of Mobility Enhancement in Short-Channel Ultra-thin Body Double-Gate MOS-

FETs," *Proceedings of the 36th European Solid-State Device Research Conference (ESSDERC)*, pp. 367–370, 2006.

[46] T. Skotnicki, G. Meckel, and T. Pedron, "The Voltage-Doping Transformation: A New Approach to the Modeling of MOSFET Short-Channel Effects," *IEEE Electron Device Letters*, vol. 9, no. 3, pp. 109–112, March 1988.

[47] T. Skotnicki, "Heading for decananometer CMOS - Is navigation among icebergs still a viable strategy?" *Proceedings of the 30th European Solid-State Device Research Conference (ESSDERC)*, pp. 19–33, September 2000.

[48] F. Balestra, S. Cristoloveanu, M. Bénachir, J. Brini, and T. Elewa, "Double-Gate Silicon on Insulator Transistor with Volume Inversion: A New Device with Greatly Enhanced Performance," *IEEE Electron Device Letters*, vol. 8, pp. 410–412, 1987.

[49] F. Gámiz and M. V. Fischetti, "Monte Carlo simulation of double-gate silicon-on-insulator inversion layers: The role of volume inversion," *Journal of Applied Physics*, vol. 89, no. 10, pp. 5478–5487, May 2001.

[50] G. A. Kathawala, B. Winstead, and U. Ravaioli, "Monte Carlo Simulations of Double-Gate MOSFETs," *IEEE Transactions on Electron Devices*, vol. 50, no. 12, pp. 2467–2473, December 2003.

[51] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Reviews of Modern Physics*, vol. 54, no. 2, pp. 437–672, April 1982.

[52] T. Ernst, S. Cristoloveanu, G. Ghibaudo, T. Ouisse, S. Horiguchi, Y. Ono, Y. Takahashi, and K. Murase, "Ultimately Thin Double-Gate SOI MOSFETs," *IEEE Transactions on Electron Devices*, vol. 50, no. 3, pp. 830–838, March 2003.

[53] F. Gámiz, J. B. Roldán, and J. A. López-Villanueva, "Phonon-limited electron mobility in ultrathin silicon-on-insulator inversion layers," *Journal of Applied Physics*, vol. 83, no. 9, pp. 4802–4806, 1 May 1998.

[54] L. Donetti, F. Gámiz, J. B. Roldán, and A. Godoy, "Confined acoustic phonons in ultrathin SOI layers," *Journal of Computational Electronics*, vol. 5, pp. 199–203, 2006.

[55] M. Shoji and S. Horiguchi, "Phonon-limited inversion layer electron mobility in extremely thin Si layer of silicon-on-insulator metal-oxide-semiconductor field-effect transistor," *Journal of Applied Physics*, vol. 82, no. 12, p. 6096, December 1997.

[56] ——, "Electronic structures and phonon-limited electron mobility of double-gate silicon-on-insulator Si inversion layers," *Journal of Applied Physics*, vol. 85, no. 5, pp. 2722–2731, March 1999.

[57] L. Donetti, F. Gámiz, J. B. Roldán, and A. Godoy, "Acoustic phonon confinement in silicon nanolayers: Effect on electron mobility," *Journal of Applied Physics*, vol. 100, p. 013701, 2006.

[58] C. M. S. Torres, A. Zwick, F. Poinsotte, J. Groenen, M. Prunnila, J. Ahopelto, A. Mlayah, and V. Paillard, "Observations of confined acoustic phonons in silicon membranes," *Physica Status Solidi (c)*, vol. 1, no. 11, pp. 2609–2612, 2004.

[59] L. Donetti, F. Gámiz, N. Rodriguez, F. Jimenez, and C. Sampedro, "Influence of acoustic phonon confinement on electron mobility in ultrathin silicon on insulator layers," *Applied Physics Letters*, vol. 88, p. 122108, 2006.

[60] G. Ferrari, J. R. Watling, S. Roy, J. R. Barker, and A. Asenov, "Beyond SiO_2 technology: Simulation of the impact of high-κ dielectrics on mobility," *Journal of Non-Crystaline Solids*, vol. 353, pp. 630–634, 2007.

[61] S. Takagi, J. Koga, and A. Toriumi, “Subband Structure Engineering for Performance Enhancement of Si MOSFETs,” *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 219–222, 7-10 December 1997.

[62] D. Vasileska and S. S. Ahmed, “Narrow-Width SOI Devices: The Role of Quantum-Mechanical Size Quantization Effect and Unintentional Doping on the Device Operation,” *IEEE Transactions on Electron Devices*, vol. 52, no. 2, pp. 227–236, February 2005.

[63] D. Vasileska, H. R. Khan, S. S. Ahmed, C. Ringhofer, and C. Heitzinger, “Quantum and Coulomb Effects in Nanodevices,” *International Journal of Nanoscience*, vol. 4, no. 3, pp. 305–361, 2005.

[64] M. V. Fischetti and S. E. Laux, “Monte Carlo study of electron transport in silicon inversion layers,” *Physical Review B*, vol. 48, no. 4, pp. 2244–2274, July 1993.

[65] S. Jallepalli, J. Bude, W.-K. Shih, M. R. Pinto, C. M. Maziar, and J. A. F. Tasch, “Electron and Hole Quantization and Their Impact on Deep Submicron Silicon p- and n-MOSFET Characteristics,” *IEEE Transactions on Electron Devices*, vol. 44, no. 2, pp. 297–303, 1997.

[66] K. Uchida, H. Watanabe, A. Kinoshita, J. Koga, T. Numata, and S. Takagi, “Experimental Study on Carrier Transport Mechanism in Ultrathin-body SOI n- and p-MOSFETs with SOI Thickness less than 5 nm,” *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 47–50, 2002.

[67] S. Takagi, J. Koga, and A. Toriumi, “Mobility Enhancement of SOI MOSFETs due to Subband Modulation in Ultrathin SOI Films,” *Japanese Journal of Applied Physics*, vol. 37, no. 3B, pp. 1289–1294, March 1998.

[68] F. Gámiz, "Temperature behaviour of electron mobility in double-gate silicon on insulator transistors," *Semiconductor Science and Technology*, vol. 19, pp. 113–119, 2004.

[69] F. Gámiz, F. Jiménez-Molinos, J. B. Roldán, and P. Cartujo-Cassinello, "Coulomb scattering model for ultrathin silicon-on-insulator inversion layers," *Applied Physics Letters*, vol. 80, no. 20, pp. 3835–3837, May 2002.

[70] J. Koga, S. Takagi, and A. Toriumi, "Influences of Buried-Oxide Interface on Inverson-Layer Mobility in Ultra-Thin SOI MOSFETs," *IEEE Transactions on Electron Devices*, vol. 49, no. 6, pp. 1042–1048, June 2002.

[71] K. Uchida, J. Koga, and S. Takagi, "Experimental Study on Carrier Transport Mechanisms in Double- and Single-Gate Ultrathin-body MOSFETs -Coulomb Scattering, Volume Inversion, and δT_{SOI}-Induced Scattering-," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 33.5.1–33.5.4, 8-10 December 2003.

[72] F. M. Bufler, A. Schenk, and W. Fichtner, "Monte Carlo, Hydrodynamic and Drift-Diffusion Simulation of Scaled Double-Gate MOSFETs," *Journal of Computational Electronics*, vol. 2, no. 2, pp. 81–84, December 2003.

[73] E. Sangiorgi and M. R. Pinto, "A Semi-Empirical Model of Surface Scattering for Monte Carlo Simulation of Silicon n-MOSFET's," *IEEE Transactions on Electron Devices*, vol. 39, no. 2, pp. 356–361, February 1992.

[74] 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. 1764–1771, 1 February 2001.

[75] B. Majikusiak, T. Janik, and J. Walczak, "Semiconductor Thickness Effects in the Double-Gate SOI MOSFET," *IEEE Transactions on Electron Devices*, vol. 45, no. 5, pp. 1127–1134, 1998.

[76] F. Gámiz, J. B. Roldán, J. A. López-Villanueva, P. Cartujo-Cassinello, and J. E. Carceller, "Surface roughness at the Si-SiO_2 interfaces in fully depleted silicon-on-insulator inversion layers," *Journal of Applied Physics*, vol. 86, no. 12, pp. 6854–6863, December 1999.

[77] 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," *IEEE Transactions on Electron Devices*, vol. 50, no. 12, pp. 2445–2455, December 2003.

[78] A. Gold, "Electronic transport properties of a two-dimensional electron gas in a silicon quantum-well structure at low temperature," *Physical Review B*, vol. 35, no. 2, pp. 723–733, January 1987.

[79] H. Sakaki, T. Noda, K. Hirakawa, M. Tanaka, and T. Matsusue, "Interface roughness scattering in GaAs/AIAs quantum wells," *Applied Physics Letters*, vol. 51, no. 23, pp. 1934–1936, December 1987.

[80] S. Venkatesan, G. W. Neudeck, and R. F. Pierret, "Dual-Gate Operation and Volume Inversion in n-Channel SOI MOSFET's," *IEEE Electron Device Letters*, vol. 12, no. 1, pp. 44–46, 1992.

[81] T. Grasser, A. Gehring, and S. Selberherr, "Recent Advances in Transport Modeling for Miniaturized CMOS Devices," *Fourth IEEE International Caracas Conference on Devices, Circuits and Systems*, pp. D027–1 – D027–8, 2002.

[82] S. E. Laux and M. V. Fischetti, "Issues in Modeling Small Devices," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 523–526, 1999.

[83] B. Winstead and U. Ravaioli, “A Quantum Correction Based on Schrodinger Equation Applied to Monte Carlo Device Simulation,” *IEEE Transactions on Electron Devices*, vol. 50, no. 2, pp. 440–446, February 2003.

[84] H. Tsuchiya and U. Ravaioli, “Particle Monte Carlo simulation of quantum phenomena in semiconductor nanostructures,” *Journal of Applied Physics*, vol. 89, no. 7, pp. 4023–4029, April 2001.

[85] B. Wu and T. wei Tang, “Quantum Corrected Boltzmann Transport Model for Tunneling Effects,” *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 279–282, 3-5 Sept 2003.

[86] D. K. Ferry, R. Akis, and D. Vasileska, “Quantum Effects in MOSFETs: Use of an Effective Potential in 3D Monte Carlo Simulation of Ultra-Short Channel Devices,” *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 287–290, 10-13 December 2000.

[87] S. Selberherr, *Analysis and Simulation of Semiconductor Devices.* Springer-Verlag Wien, New York, 1984.

[88] C. M. Snowden, “Semiconductor device modelling,” *Reports on Progress in Physics*, vol. 48, pp. 223–275, 1985.

[89] W. L. Engl, H. K. Dirks, and B. Meinerzhagen, “Device Modeling,” *Proceedings of the IEEE*, vol. 71, no. 1, pp. 10–33, January 1983.

[90] V. Sverdlov, H. Kosina, and S. Selberherr, “Current Transport in Nanoelectronic Semiconductor Devices,” *2006 IEEE Conference on Emerging Technologies - Nanoelectronics*, pp. 490–495, January 2006.

[91] M. Lundstrom, “Assumptions and Trade-Offs in Device Simulation Programs,” *IEEE 1992 Bipolar Circuits and Technology Meeting*, pp. 35–41, 1992.

[92] A. Asenov, S. Kaya, and J. H. Davies, "Intrinsic Threshold Voltage Fluctuations in Decanano MOSFETs Due to Local Oxide Thickness Variations," *IEEE Transactions on Electron Devices*, vol. 49, no. 1, pp. 112–119, January 2002.

[93] A. Asenov, S. Kaya, and A. R. Brown, "Intrinsic Parameter Fluctuations in Decananometer MOSFETs Introduced by Gate Line Edge Roughness," *IEEE Transactions on Electron Devices*, vol. 50, no. 5, pp. 1254–1260, May 2003.

[94] J.-H. Rhew, Z. Ren, and M. S. Lundstrom, "Benchmarking Macroscopic Transport Models for Nanotransistor TCAD," *Journal of Computational Electronics*, vol. 1, pp. 385–388, 2002.

[95] ——, "A numerical study of ballistic transport in a nanoscale MOSFET," *Solid-State Electronics*, vol. 46, pp. 1899–1906, 2002.

[96] R. Granzner, V. M. Polyakov, F. Schwierz, M. Kittler, and T. Doll, "On the suitability of DD and HD models for the simulation of nanometer double-gate MOSFETs," *Physica E*, vol. 19, pp. 33–38, 2003.

[97] C. L. Alexander, G. Roy, and A. Asenov, "Random Impurity Scattering Induced Variability in Conventional Nano-Scaled MOSFETs: Ab initio Impurity Scattering Monte Carlo Simulation Study," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 1–4, 2006.

[98] B. Meinerzhagen and W. L. Engl, "The Influence of the Thermal Equilibrium Approximation on the Accuracy of Classical Two-Dimensional Numerical Modeling of Silicon Submicrometer MOS Transistors," *IEEE Transactions on Electron Devices*, vol. 35, no. 5, pp. 689–697, May 1988.

[99] J. G. Ruch, "Electron Dynamics in Short Channel Field-Effect Transistors," *IEEE Transactions on Electron Devices*, vol. 19, no. 5, pp. 652–654, May 1972.

[100] S. Y. Chou, D. A. Antoniadis, and H. I. Smith, "Observation of Electron Velocity Overshoot in Sub-100-nm-channel MOSFET's in Silicon," *IEEE Electron Device Letters*, vol. 6, no. 12, pp. 665–667, December 1985.

[101] M. Lundstrom and S. Datta, "Physical Device Simulation in a Shrinking World," *IEEE Circuits And Devices Magazine*, vol. 6, pp. 32–37, 1990.

[102] K. Banoo and M. S. Lundstrom, "Electron transport in a model Si transistor," *Solid-State Electronics*, vol. 44, pp. 1689–1695, 2000.

[103] G. A. S. Halasz, M. R. Wordeman, D. P. Kern, S. Rishton, and E. Ganin, "High Transconductance and Velocity Overshoot in NMOS Devices at the 0.1-μm Gate-Length Level," *IEEE Electron Device Letters*, vol. 9, no. 9, pp. 464–466, September 1988.

[104] S. E. Laux and M. V. Fischetti, "Monte Carlo Study of Velocity Overshoot in Switching a 0.1-Micron CMOS Inverter," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 877–880, 1997.

[105] J. D. Bude, "MOSFET Modeling Into the Ballistic Regime," *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 23–26, September 2000.

[106] M. Lundstrom, *Fundamentals of Carrier Transport (2nd Edition)*. Cambridge University Press, Cambridge, 2000.

[107] F. O. Heinz, F. M. Bufler, A. Schenk, and W. Fichtner, "Quantum transport phenomena and their modeling," in *Symposium on Nano Device Technology*, 2004, pp. 2–8.

[108] M. G. Ancona and G. J. Iafrate, "Quantum correction to the equation of state of an electron gas in a semiconductor," *Physical Review B*, vol. 39, no. 13, pp. 9536–9540, May 1989.

[109] D. Bohm, "A Suggested Interpretation of the Quantum Theory In Terms of "Hidden" Variables. I," *Physical Review*, vol. 85, no. 2, pp. 166–179, January 1952.

[110] J. R. Watling, A. R. Brown, A. Asenov, A. Svizhenko, and M. P. Anantram, "Simulation of direct source-to-drain tunnelling using the density gradient formalism: Non-equilibrium Green's function calibration," *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 267–270, 2002.

[111] M. G. Ancona, "Equations of State for Silicon Inversion Layers," *IEEE Transactions on Electron Devices*, vol. 47, no. 7, pp. 1449–1456, July 2000.

[112] C. S. Rafferty, B. Biegel, Z. Yu, M. G. Ancona, J. Bude, and R. W. Dutton, "Multi-dimensional quantum effect simulation using a density-gradient model and script-level programming techniques," *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 137–140, September 2-4 1998.

[113] A. Asenov, J. R. Watling, A. R. Brown, and D. K. Ferry, "The Use of Quantum Potentials for Confinement and Tunnelling in Semiconductor Devices," *Journal of Computational Electronics*, vol. 1, no. 4, pp. 503–513, 2002.

[114] J. R. Watling, A. R. Brown, and A. Asenov, "Can the Denisty Gradient Approach Describe the Source-Drain Tunnelling in Decanano Double-Gate MOSFETs?" *Journal of Computational Electronics*, vol. 1, pp. 289–293, 2002.

[115] A. R. Brown, F. Adamu-Lema, and A. Asenov, "Intrinsic parameter fluctuations in nanometer scale thin-body SOI devices introduced by interface roughness," *Superlattices and Microstructures*, vol. 34, pp. 283–291, 2003.

[116] D. K. Ferry, "The onset of quantization in ultra-submicron semiconductor devices," *Superlattices and Microstructures*, vol. 27, no. 2-3, pp. 61–66, 2000.

[117] D. K. Ferry, S. Ramey, L. Shifren, and R. Akis, "The Effective Potential in Device Modeling: The Good, the Bad and the Ugly," *Journal of Computational Electronics*, vol. 1, no. 1/2, pp. 59–65, July 2002.

[118] Y. Li, T. wei Tang, and X. Wang, "Modeling of Quantum Effects for Ultrathin Oxide MOS Structures With an Effective Potential," *IEEE Transactions on Nanotechnology*, vol. 1, no. 4, pp. 238–242, December 2002.

[119] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals.* McGraw-Hill Companies, New York, 1965.

[120] 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, no. 3, pp. 645–705, July 1983.

[121] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation.* Springer-Verlag Wien, New York, 1989.

[122] C. Moglestue, *Monte Carlo Simulations of Semiconductor Devices.* Chapman and Hall, 1993.

[123] K. Tomizawa, *Numerical Simulation of Submicron Semiconductor Devices.* Artech House, 1993.

[124] J. M. Higman, K. Hess, C. G. Hwang, and R. W. Dutton, "Coupled Monte Carlo-Drift Diffusion Analysis of Hot-Electron Effects in MOSFETs," *IEEE Transactions on Electron Devices*, vol. 36, no. 5, pp. 930–937, 1989.

[125] F. M. Bufler, A. Schenk, and W. Fichtner, "Efficient Monte Carlo Device Modeling," *IEEE Transactions on Electron Devices*, vol. 47, no. 10, pp. 1891–1897, October 2000.

[126] C. Jungemann and B. Meinerzhagen, “On the applicability of nonself-consistent Monte Carlo device simulations,” *IEEE Transactions on Electron Devices*, vol. 49, no. 6, pp. 1072–1074, June 2002.

[127] F. Venturi, R. K. Smith, E. C. Sangiorgi, M. R. Pinto, and B. Riccò, “A General Purpose Device Simulator Coupling Poisson and Monte Carlo Transport with Applications to Deep Submicron MOSFET’s,” *IEEE Transactions on Computer-Aided Design*, vol. 8, no. 4, pp. 360–369, April 1989.

[128] C. Moglestue, “Self-consistent Monte Carlo particle modelling of small semiconductor elements,” *Reports on Progress in Physics*, vol. 53, pp. 1333–1353, 1990.

[129] J. A. Greenfield and R. W. Dutton, “Nonplanar VLSI Device Analysis Using the Solution of Poisson’s Equation,” *IEEE Transactions on Electron Devices*, vol. ED-27, no. 8, pp. 1520–1532, August 1980.

[130] H. A. van der Vorst, “Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems,” *SIAM Journal of Scientific Computing*, vol. 13, no. 2, pp. 631–644, March 1992.

[131] U. Trottenberg, C. Oosterlee, and A. Schuller, *Multgrid.* Academic Press, San Diego, CA, 2001.

[132] C. Millar, A. Asenov, and J. R. Watling, “Excessive Over-Relaxation Method for Multigrid Poisson Solvers,” *Journal of Computational Electronics*, vol. 1, pp. 341–345, 2002.

[133] M. V. Fischetti and S. E. Laux, “Monte Carlo analysis of electron transport in small semiconductor devicews including band-structure and space-charge effects,” *Physical Review B*, vol. 38, no. 14, pp. 9721–9745, November 1988.

[134] C. Jungemann, S. Keith, M. Bartels, and B. Meinerzhagen, "Efficient Full-Band Monte Carlo Simulation of Silicon Devices," *IEICE Transactions on Electronics*, vol. E82-C, no. 6, June 1999.

[135] A.-B. Chen and A. Sher, "Electronic structure of III-V semiconductors and alloys using simple orbitals," *Physical Review B*, vol. 22, no. 8, pp. 3886–3896, October 1980.

[136] C. M. Goringe, D. R. Bowler, and E. Hernàndez, "Tight-binding modelling of materials," *Reports on Progress in Physics*, vol. 60, pp. 1447–1512, 1997.

[137] B. Ghosh, X. Wang, X.-F. Fan, L. F. Register, and S. K. Banerjee, "Monte Carlo Study of Germanium n- and pMOSFETs," *IEEE Transactions on Electron Devices*, vol. 52, no. 4, pp. 547–553, April 2005.

[138] M. V. Fischetti, S. E. Laux, and E. Crabbé, "Understanding hot-electron transport in silicon devices: Is there a shortcut?" *Journal of Applied Physics*, vol. 78, no. 2, pp. 1058–1087, July 1995.

[139] 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, January 1996.

[140] S. C. Williams, K. W. Kim, and W. C. Holton, "Ensemble Monte Carlo Study of Channel Quantization in a 25-nm n-MOSFET," *IEEE Transactions on Electron Devices*, vol. 47, no. 10, pp. 1864–1872, October 2000.

[141] U. Ravaioli, B. Winstead, C. Wordelman, and A. Kepkep, "Monte Carlo simulation for ultra-small MOS devices," *Superlattices and Microstructures*, vol. 27, no. 2/3, pp. 137–145, 2000.

[142] C. Alexander, J. R. Watling, A. R. Brown, and A. Asenov, "Artificial carrier heating due to the introduction of ab initio Coulomb scattering

in Monte Carlo simulations," *Superlattices and Microstructures*, vol. 34, pp. 319–326, 2003.

[143] C. L. Alexander, A. R. Brown, J. R. Watling, and A. Asenov, "Impact of Single Charge Trapping in Nano-MOSFETs - Electrostatics Versus Transport Effects," *IEEE Transactions on Nanotechnology*, vol. 4, no. 3, pp. 339–344, May 2005.

[144] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles*. Institue of Physics Publishing, London, 1988.

[145] C. Alexander, J. R. Watling, and A. Asenov, "Numerical carrier heating whem implementing P^3M to study small volume variations," *Semiconductor Science and Technology*, vol. 19, pp. S139–S141, 2004.

[146] C. J. Wordelman and U. Ravaioli, "Intergration of a Particle-Particle-Particle-Mesh Algorithm with the Ensemble Monte Carlo Method for the Simulation of Ultra-Small Semiconductor Devices," *IEEE Transactions on Electron Devices*, vol. 47, no. 2, pp. 410–416, February 2000.

[147] F. Gámiz, J. B. Roldán, A. Godoy, J. E. Carceller, and P. Cartujo, "Double gate silicon on insulator transistors. A Monte Carlo study," *Solid-State Electronics*, vol. 48, pp. 937–945, 2004.

[148] R. Ravishankar, G. Kathawala, and U. Ravaioli, "Comparison of Monte Carlo and NEGF Simulations of Double Gate MOSFETs," *Journal of Computational Electronics*, vol. 4, pp. 39–43, 2005.

[149] B. Winstead, H. Tsuchiya, and U. Ravaioli, "Comparison of Quantum Corrections for Monte Carlo Simulations," *Journal of Computational Electronics*, vol. 1, no. 1, pp. 201–207, 2002.

[150] H. Tsuchiya and T. Miyoshi, "Quantum mechanical Monte Carlo approach to electron transport at heterointerface," *Superlattices and Microstructures*, vol. 27, no. 5-6, pp. 529–532, 2000.

[151] H. Takeda, N. Mori, and C. Hamaguchi, "Quantum Effects on Transport Characteristics in Ultra-Small MOSFETs," *Journal of Computational Electronics*, vol. 2, pp. 119–122, 2003.

[152] L. Lucci, P. Palestri, D. Esseni, and L. Selmi, "Multi-subband Monte Carlo modeling of nano-MOSFETs with strong vertical quantization and electron gas degeneration," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 631–634, 2005.

[153] J. Saint-Martin, D. Querlioz, A. Bournel, and P. Dollfus, "Efficient multi sub-band Monte Carlo simulation of nano-scaled Double Gate MOSFETs," *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 216–219, 2006.

[154] H. Tsuchiya and T. Miyoshi, "Quantum Transport Modeling of Ultrasmall Semiconductor Devices," *IEICE Transactions on Electronics*, vol. E82-C, no. 6, pp. 880–888, June 1999.

[155] H. Tsuchiya, M. Horino, M. Ogawa, and T. Miyoshi, "Quantum Transport Simulation of Ultrathin and Ultrashort Silicon-On-Insulator Metal-Oxide-Semiconductor Field-Effect Transistors," *Japanese Journal of Applied Physics*, vol. 42, no. 12, pp. 7238–7243, December 2003.

[156] M. Ogawa, H. Tsuchiya, and T. Miyoshi, "Quantum Transport Modeling in Nano-Scale Devices," *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 261–266, 2002.

[157] H. Tsuchiya, M. Horino, and T. Miyoshi, "Quantum Monte Carlo Device Simulation of Nano-Scaled SOI-MOSFETs," *Journal of Computational Electronics*, vol. 2, pp. 91–95, 2003.

[158] H. Tsuchiya, A. Svizhenko, M. P. Anantram, M. Ogawa, and T. Miyoshi, "Comparison of Non-Equilibrium Green's Function and Quantum-Corrected Monte Carlo Approaches in Nano MOS Simulation," *Journal of Computational Electronics*, vol. 4, pp. 35–38, 2005.

[159] T. wei Tang and B. Wu, "Quantum Corrected Monte Carlo Simulation of Semiconductor Devices Using the Effective Conduction-Band Edge Method," *Journal of Computational Electronics*, vol. 2, no. 2-4, pp. 131–135, 2003.

[160] B. Wu and T. wei Tang, "The Effective Conduction Band Edge Method of Quantum Correction to the Monte Carlo Device Simulation," *Journal of Computational Electronics*, vol. 3, no. 3/4, pp. 347–350, October 2004.

[161] C. Sampedro-Matarin, F. Gámiz, A. Godoy, and F. J. G. Ruiz, "The Multivalley Effective Conduction Band-Edge Method for Monte Carlo Simulation of Nanoscale Structures," *IEEE Transactions on Electron Devices*, vol. 53, no. 11, pp. 2703–2710, November 2006.

[162] C. Sampedro, F. Gámiz, A. Godoy, and F. G. Ruiz, "Quantum corrected Ensemble Monte Carlo simulation of UTB-DGSOI Contribution of Volume Inversion and Inter-Subband Modulation effects," *2006 IEEE International SOI Conference Proceedings*, pp. 79–80, 2006.

[163] C. Sampedro, F. Gámiz, A. Godoy, and F. Jiménez-Molinos, "Quantum Ensemble Monte Carlo simulation of silicon-based nanodevices," *Journal of Computational Electronics*, vol. 6, no. 1-3, pp. 41–44, 2007.

[164] K. Kalna and A. Asenov, "Quantum Corrections in the Monte Carlo Simulations of Scaled PHEMTs with Multiple Delta Doping," *Journal of Computational Electronics*, vol. 1, pp. 257–261, 2002.

[165] S. M. Ramey and D. K. Ferry, "Implementation of Surface Roughness Scattering in Monte Carlo Modeling of Thin SOI MOSFETs Using the Effective Potential," *IEEE Transactions on Nanotechnology*, vol. 2, no. 2, pp. 110–114, June 2003.

[166] P. Palestri, S. Eminente, D. Esseni, C. Fiegna, E. Sangiorgi, and L. Selmi, "An improved semi-classical Monte-Carlo approach for nano-scale MOSFET simulation," *Solid-State Electronics*, vol. 49, pp. 727–732, 2005.

[167] S. M. Ramey and D. K. Ferry, "3D Monte Carlo Modeling of Thin SOI MOSFETs Including the Effective Potential and Random Dopant Distribution," *Journal of Computational Electronics*, vol. 1, pp. 267–271, 2002.

[168] M.-A. Jaud, S. Barraud, P. Dollfus, H. Jaouen, F. D. Crecy, and G. L. Carval, "Validity of the effective potential approach for the simulation of quantum confinement effects: A Monte-Carlo study," *Journal of Computational Electronics*, vol. 5, no. 2-3, pp. 171–175, 2006.

[169] M.-A. Jaud, S. Barraud, P. Dollfus, H. Jaouen, and G. L. Carval, "Monte Carlo simulation of ultimate DGMOS based on a Pearson Effective Potential formalism," *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 196–199, 2006.

[170] ——, "Pearson versus gaussian effective potentials for quantum-corrected Monte-Carlo simulation," *Journal of Computational Electronics*, vol. 6, no. 1-3, pp. 19–22, 2007.

[171] X.-F. Fan, X. Wang, B. Winstead, L. F. Register, U. Ravaioli, and S. K. Banerjee, "MC Simulation of Straind-Si MOSFET With Full-Band Structure and Quantum Correction," *IEEE Transactions on Electron Devices*, vol. 51, no. 6, pp. 962–970, June 2004.

[172] G. A. Kathawala and U. Ravaioli, "3-D Monte Carlo simulations of FinFETs," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 29.2.1–29.2.4, December 2003.

[173] S. Datta, "Nanoscale device modeling: the Green's function method," *Superlattices and Microstructures*, vol. 28, no. 4, pp. 253–278, 2000.

[174] ——, "The Non-Equilibrium Green's Function (NEGF) Formalism: An Elementary Introduction," *International Electron Devices Meeting (IEDM) Technical Digest*, no. 703-706, 2002.

[175] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal, "Two-dimensional quantum mechanical modeling of nanotransistors," *Journal of Applied Physics*, vol. 91, no. 4, pp. 2343–2354, February 2002.

[176] S. Datta, *Electronic Transport in Mesoscopic Systems*, ser. Cambudge Studies in Semiconductor Physics and Mircoelectronic Engineering. Cambridge University Press, Cambridge, 1997.

[177] R. Venugopal, M. Paulsson, S. Goasguen, S. Datta, and M. S. Lundstrom, "A simple quantum mechanical treatment of scattering in nanoscale transistors," *Journal of Applied Physics*, vol. 93, no. 9, pp. 5613–5625, May 2003.

[178] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom, "nanoMOS 2.5: A Two-Dimensional Simulator for Quantum Transport in Double-Gate MOSFETs," *IEEE Transactions on Electron Devices*, vol. 50, no. 9, pp. 1914–1925, September 2003.

[179] S. Hasan, J. Wang, and M. Lundstrom, "Device design and manufacturing issues for 10nm-scale MOSFETs: a computational study," *Solid-State Electronics*, vol. 48, pp. 867–875, 2004.

[180] R. Lake, G. Klimeck, R. C. Bown, and D. Jovanovic, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *Journal of Applied Physics*, vol. 81, no. 12, pp. 7845–7869, June 1997.

[181] A. Svizhenko and M. P. Anantram, "Role of Scattering in Nanotransistors," *IEEE Transactions on Electron Devices*, vol. 50, no. 6, pp. 1459–1466, June 2003.

[182] X. Shao and Z. Yu, "Nanoscale FinFET simulation: A quasi-3D quantum mechanical model using NEGF," *Solid-State Electronics*, vol. 49, pp. 1435–1445, 2005.

[183] A. Martinez, J. R. Barker, A. Svizhenko, M. P. Anantram, A. R. Brown, B. Biegel, and A. Asenov, "The impact of unintentional discrete charges in a nominally undoped channel of a thin body double gate MOSFET: classical to full quantum simulation," *Journal of Physics: Conference Series*, vol. 38, pp. 192–195, 2006.

[184] J. Fonseca and S. Kaya, "Accurate treatment of interface roughness in nanoscale DG MOSFETs using non-equilibrium Green's functions," *Solid-State Electronics*, vol. 48, pp. 1843–1847, 2004.

[185] A. Martinez, A. Svizhenko, M. P. Anantram, J. R. Barker, A. R. Brown, and A. Asenov, "A study of the effect of the interface roughness on a DG-MOSFET using a full 2D NEGF technique," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 627–630, 2005.

[186] A. Martinez, A. Svizhenko, M. P. Anantram, J. R. Barker, and A. Asenov, "A NEGF Study of the Effect of Surface Roughness on CMOS Nanotransistors," *Journal of Physics: Conference Series*, vol. 35, pp. 269–274, 2006.

[187] A. Martinez, J. R. Barker, A. Asenov, A. Svizhenko, and M. P. Anantram, "Developing a full 3D NEGF simulator with random dopant and interface roughness," *Journal of Computational Electronics*, vol. 6, no. 1-3, pp. 215–218, 2007.

[188] A. Asenov, A. R. Brown, J. H. Davies, and S. Saini, "Hierarchical Approach to "Atomistic" 3-D MOSFET Simulation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 18, no. 11, pp. 1558–1565, November 1999.

[189] A. R. Brown, A. Asenov, and J. R. Watling, “Intrinsic Fluctuations in Sub 10-nm Double-Gate MOSFETs Introduced by Discreteness of Charge and Matter,” *IEEE Transactions on Nanotechnology*, vol. 1, no. 4, pp. 195–200, December 2002.

[190] S. Jin, Y. J. Park, and H. S. Min, “Simulation of Quantum Effects in the Nano-scale Semiconductor Device,” *Journal of Semiconductor Technology and Science*, vol. 4, no. 1, pp. 32–40, March 2004.

[191] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, “Surface roughness at the Si(100)-SiO_2 interface,” *Physical Review B*, vol. 32, no. 12, pp. 8171–8189, 15 December 1985.

[192] T. Yoshinobu, A. Iwamoto, and H. Iwasaki, “Scaling Analysis of SiO_2/Si Interface Roughness by Atomic Force Microscopy,” *Japanese Journal of Applied Physics*, vol. 33, pp. 383–387, January 1994.

[193] D. Vasileska and D. K. Ferry, “Scaled Silicon MOSFET’s: Universal Mobility Behavior,” *IEEE Transactions on Electron Devices*, vol. 44, no. 4, pp. 577–583, April 1997.

[194] D. Esseni, “On the Modelling of Surface Roughness Limited Mobility in SOI MOSFETs and Its Correlation to the Transistor Effective Field,” *IEEE Transactions on Electron Devices*, vol. 51, no. 3, pp. 394–401, March 2004.

[195] B. Sell, D. Schumann, and W. H. Krautschneider, “Fast Interface Characterization of Tunnel Oxide MOS Structures,” *IEEE Transactions on Nanotechnology*, vol. 1, no. 2, pp. 110–113, June 2002.

[196] S. E. Laux, “On Particle-Mesh Coupling in Monte Carlo Semiconductor Device Simulation,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, no. 10, pp. 1266–1277, October 1996.

[197] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters," *Journal of Chemical Physics*, vol. 76, no. 1, pp. 637–649, January 1982.

[198] P. W. Rambo and J. Denavit, "Time Stability of Monte Carlo Device Simulation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 12, no. 11, pp. 1734–1741, November 1993.

[199] P. Palestri, N. Barin, D. Esseni, and C. Fiegna, "Stabilty of Self-Consistent Monte Carlo Simulations: Effects of the Grid Size and of the Coupling Scheme," *IEEE Transactions on Electron Devices*, vol. 53, no. 6, pp. 1433–1442, June 2006.

[200] D. L. Woolard, H. Tian, M. A. Littlejohn, and K. W. Kim, "Efficient Ohmic Boundary Conditions for the Monte Carlo Simulation of Electron Transport," *IEEE Transactions on Electron Devices*, vol. 41, no. 4, pp. 601–606, April 1994.

[201] T. González and D. Pardo, "Physical Models of Ohmic Contact for Monte Carlo Device Simulation," *Solid-State Electronics*, vol. 39, no. 4, pp. 555–562, 1996.

[202] I. Riolino, M. Braccioli, L. Lucci, D. Esseni, C. Fiegna, P. Palestri, and L. Selmi, "Monte-Carlo Simulation of Decananometric Double-Gate SOI devices: Multi-Subband vs. 3D-Electron Gas with Quantum Corrections," *Proceedings of the 36th European Solid-State Device Research Conference (ESSDERC)*, pp. 162–165, 2006.

[203] M. G. Ancona, D. Yergeau, Z. Yu, and B. A. Biegel, "On Ohmic Boundary Conditions for Density-Gradient Theory," *Journal of Computational Electronics*, vol. 1, pp. 103–107, 2002.

[204] R. Clerc, P. Palestri, and A. Abramo, "Invesigation on Convergence and Stability of Self-Consistent Monte Carlo Device Simulations," *Proceedings of the 32nd European Solid-State Device Research Conference (ESSDERC)*, pp. 191–194, 2002.