



<https://theses.gla.ac.uk/>

Theses Digitisation:

<https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/>

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk

***Ab initio* Scattering From Random Discrete
Charges and its Impact on the Intrinsic
Parameter Fluctuations in Nano-CMOS Devices**

Craig L. Alexander

July 2005

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the
University of Glasgow,

Department of Electronics and Electrical Engineering.

©Craig L. Alexander April 2005. No quotation from this thesis, nor any information
derived therefrom, may be published without the author's prior written consent.

ProQuest Number: 10800628

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10800628

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

GLASGOW
UNIVERSITY
LIBRARY:

Abstract

This thesis is concerned with the Monte Carlo simulation of device parameter variation associated with the discrete nature and random variation of ionized impurity atoms within ultra-small conventional n -MOS devices. In particular, the Monte Carlo method is applied to accurately resolve electron interactions with individual ionized impurity atoms and in so doing capture the variation in impurity scattering associated with randomly configured dopant distributions. To date, variation in transport due to position dependent variation in Coulomb scattering has not received any attention although it is expected to increase the inherent device parameter variation.

A detailed methodology for the accurate treatment of Coulomb scattering within the Ensemble Monte Carlo framework is presented and verified. Improvement over existing methodologies is presented with a short-range force model that significantly reduces errors in conservation of energy during short-range attractive interactions compared with models proposed in similar work. Details of the simulated reproduction of bulk mobility are thoroughly presented to validate the method, while to date such detail is not to be found anywhere in the literature.

A charge assignment method is developed to be applied to traditional 'continuously' doped regions in order to allow a consistent description of doping charge when combined with 'atomistic' doping assigned via the Cloud-In-Cell scheme. The charge assignment method also represents the only consistent description of electron charge assigned via CIC and the continuous doping charge.

Trapping of a single electron in a series of scaled n -channel MOSFETs was studied with the *ab initio* Coulomb scattering method and is consistently seen to increase the Random Telegraph Signal, associated with the trapping and de-trapping of such charges, when compared with Drift-Diffusion simulations. It is seen that the electrostatic influence

of the trapped charge is most prominent at low applied gate voltages where it accounts for nearly 70 – 80% of the total current reduction when including transport variation in devices with channel lengths of 30 – 10 nm. At high gate voltages, transport variation is the dominant factor with the electrostatic impact accounting for only 40 – 60% of the total variation in the same devices.

Extending this treatment to an ensemble of atomistic devices, it is seen that the inclusion of transport variations significantly increases the distribution in device parameters and that the transport variation is significantly dependent upon the specific dopant distribution. Within an ensemble of 50 'atomistic' devices, it was seen from Drift-Diffusion simulation that the average current showed a 3.0% increase over the continuously doped device, while Monte Carlo simulations resulted in a decrease in average current of 1.5%. The standard deviation of the current distribution from Drift-Diffusion simulations was 2.4% while, significantly, Monte Carlo simulations returned a value of 6.7%. This has implications for the published data obtained from Drift-Diffusion simulations which will underestimate the variation.

Acknowledgements

Having spent a long time writing this thesis, I realise that a great deal in the completion of this work is owed to both Prof. Asen Asenov and Dr. Jeremy Watling. Without the guidance and motivation given by my two supervisors, this thesis would stop somewhere in the middle of chapter 3. In particular, I'm grateful for their open door policy and the many, albeit often lengthy, discussions with Jeremy, without who's insight into Monte Carlo simulation I'd be at a loss. Although, he always let me do things he knew wouldn't work, so really..... He'd probably like me to say he taught me everything I knew about the subject, but that would be overlooking Karol Kalna who I'd like to thank for filling in the gaps. Similarly I'd like to thank Campbell Millar, as the only other person within the group doing 3D particle simulations, who's own bugs, useless results and wasted simulation runs always reassured me, but mostly for endless help with Multigrid. Yet more thanks to Andrew Brown who contributed to the results in the last chapter and for entertaining yet more endless questions about his code.

Sharing the last 4 years in my office, the best office, were two other PhD students and friends who I can't hold accountable enough for the times miss-spent playing games, shooting fruit and paracetamol, chatting about stuff, browsing the web, Country and Western musics etc... They are, in order as you walk in the door and as shown in figure 1, Gareth Roy and Fikru Adamu-Lema. I owe them a serious many thanks for invaluable help and for making the years more fun than work.

Nearly at an end I'd like to give special thanks to both my quizzical mum (now you can see what I do with my time if you try to read this!) and dad who always trusted me enough to let me do what I wanted and always believed I could succeed in anything. I thank my brother, Thomas, and sister, Yvonne, who harassed me endlessly but otherwise took interest on some level, I think.



Figure 1: This is Gareth (left) and Fikru (right)

Final thanks to my closest friends Graham Forbes and Andrew Jackson. Graham, whose ceaseless need for extreme activities, holidays and a house in Florida, is one of the reasons this is 6 months overdue, and Andrew, who's unstoppable good nature, appetite for some food a few beers and stupid games is another! Thanks to Xinzi Yu, the smallest, dirtiest, smuttiest and funniest export from China, who will soon be writing her own thesis, for teaching me lots of Chinese, Graham's parents, Graham and Joan, for countless free meals, generosity and hospitality above and beyond anything that is reasonable to expect from another family. Thanks also to David Cameron, Chuff Johnson, Pamela Jordan and Alan "Captain J" Forbes. Form all the guys at Capoeira I must thank fearless Stephen Bonar as constantly optimistic and a good guy to know, Karen MacKenzie for a lot of support in all matters and especially to Marie Louise Gibson for a summer I'll always remember.

THANK YOU

Contents

1	Introduction	1
1.1	Random Dopants	3
1.2	This Work	5
2	Intrinsic Fluctuations and Device Parameter Variations	8
2.1	Introduction	8
2.2	Random Dopants	11
2.2.1	Experimental Observations	12
2.2.2	Modelling Approaches	13
2.2.3	Single Charge Trapping	17
2.2.4	Limitation of Drift Diffusion Simulation	18
2.2.5	Monte Carlo Simulation	19
2.2.6	Random Dopants in Non-Conventional Si Devices	20
2.3	Line Edge Roughness and Oxide Thickness Fluctuations	21
2.3.1	Line Edge Roughness	21
2.3.2	Oxide Thickness Fluctuations	23
2.4	Additional Sources of Fluctuation	26
2.5	Conclusions	27

3	Monte Carlo Device Simulation	29
3.1	Introduction	29
3.2	Boltzmann Transport Equation	30
3.2.1	Single Particle Distribution Function	30
3.2.2	Classical BTE	31
3.2.3	Semi-Classical Equations of Motion	31
3.2.4	Semi-Classical BTE	33
3.3	Monte Carlo Solution of the BTE	34
3.3.1	Band Structure	37
3.3.2	Scattering Mechanisms and Rates	41
3.3.2.1	Acoustic phonons	42
3.3.2.2	Optical phonons	43
3.3.2.3	Ionised Impurity Scattering	44
3.3.2.4	Total scattering rate	45
3.3.3	Free Flight	46
3.3.4	Selection of Scattering Event	48
3.3.5	Statistics Gathering	49
3.4	Bulk Transport in silicon	51
3.4.1	Calibration	52
3.4.2	Bulk Velocity-Field Characteristics	53
3.4.3	Bulk Energy-Field Curve	55
3.4.4	Bulk Concentration Dependent Mobility	55
3.4.5	Transient Response	55
3.5	Device Simulation	58
3.5.1	Solution Domain and Mesh Design	58
3.5.2	Charge Assignment	60

3.5.3	Force Evaluation and Interpolation	61
3.5.4	Field Adjusting Time Step	63
3.5.5	Boundary Conditions	63
3.5.6	Current Calculation	65
3.6	Example: Self Consistent N-I-N Diode	66
3.6.1	Electrostatic potential and electron concentration	67
3.6.2	Electron energy and velocity	67
3.7	Conclusion	72
4	Ionized Impurity Scattering	75
4.1	Introduction	75
4.2	Ionized Impurity Scattering Rates	76
4.2.1	Brooks-Herring Model	79
4.2.2	Conwell-Weisskopf Model	81
4.2.3	Third Body Exclusion Model	82
4.2.4	Limitation of scattering rates	84
4.3	Integrating the Equations of Motion	86
4.3.1	Euler Integration	86
4.3.2	Runge-Kutta Schemes	87
4.3.3	Classical Dynamics	88
4.3.4	Leap-Frog	89
4.3.5	Velocity-Verlet	90
4.4	Validation of Propagation	93
4.4.1	Reflection at surface	94
4.4.2	Propagation in an orbit	95
4.4.3	Coulomb scattering	96

4.5	<i>ab initio</i> Ionized Impurity Scattering	100
4.5.1	Short Range Correction, F^{sr}	102
4.5.2	Reference Force, F^{ref}	102
4.5.3	Analytic Short Range Force, F^*	105
4.5.4	Short Range Radius, r_{sr}	112
4.5.5	Application of Force Correction	113
4.6	<i>ab initio</i> Simulation of Doping Concentration Dependent Mobility in Silicon	114
4.6.1	Bulk Simulation Structure	115
4.6.2	Bulk mobility results	118
4.7	Atomistic N-I-N Diode	126
4.7.1	'atomistic' electrostatic potential and electron concentration . . .	126
4.7.2	'atomistic' electron energy and velocity	128
4.8	Conclusion	129
5	Mobility Variations	133
5.1	Introduction	133
5.2	Continuous-Cloud-In-Cell Charge Assignment	135
5.2.1	Node Centred Mesh	136
5.2.2	Node Aligned Mesh	142
5.2.3	Application of CCIC Charge Assignment	145
5.3	Single Trapped Charge	148
5.3.1	Introduction	149
5.3.2	Simulation methodology	151
5.3.3	Results	155
5.3.3.1	Percentage Current Reduction in 30nm MOSFET . . .	155
5.3.3.2	Separation of Electrostatic & Coulomb Scattering Effects	157

5.3.3.3	Percentage Current Reduction in 20nm and 10nm MOS-FETs	164
5.3.4	Conclusion	165
5.4	Atomistic MOSFET	168
5.4.1	Introduction	169
5.4.2	Simulation Methodology	170
5.4.3	Results	171
5.4.3.1	Current distribution	171
5.4.3.2	Percentage Change in Current	172
5.4.3.3	Drift Diffusion Vs. Monte Carlo	174
5.4.3.4	Self-consistent Monte Carlo simulation	182
5.4.4	Conclusion	186
5.5	Summary	187
6	Conclusion	189
6.1	The Monte Carlo Simulator	190
6.2	<i>ab initio</i> Coulomb Scattering	190
6.3	Transport Variation	192
6.4	Future Work	194
	Bibliography	195

List of Figures

1	Gareth and Fikru	5
2.1	Ideal device Vs. realistic ultrasmall device	9
2.2	Simple model for deriving analytical V_t fluctuations	14
2.3	Atomistic device following process simulation	16
2.4	Variation in resist aggregates lead directly to LER	22
2.5	Potential fluctuations associated with combined LER and RD	24
2.6	SEM image of oxide thickness fluctuations	25
2.7	Illustration of high-k gate stack fluctuations	26
3.1	Monte Carlo flow chart	35
3.2	Si band structure	37
3.3	Illustration of analytical band model	40
3.4	Phonon dispersion approximation	43
3.5	Calculated scattering rate in silicon at 300K	45
3.6	Random selection of scattering mechanism	48
3.7	Statistics gathering in a synchronous ensemble	50
3.8	Simulated bulk velocity-field curves in silicon	53
3.9	Simulated bulk energy-field curves in silicon	54
3.10	Simulated concentration dependent bulk mobility in silicon	56

3.11	Simulation of velocity overshoot	57
3.12	Device discretisation meshing schemes	59
3.13	Illustration of simulated N-I-N device	66
3.14	3D potential distribution in N-I-N diode	68
3.15	1D averaged potential profile in N-I-N diode	68
3.16	3D electron concentration distribution in N-I-N diode	69
3.17	1D averaged electron concentration profile in N-I-N diode	69
3.18	3D electron energy distribution in N-I-N diode	70
3.19	1D averaged electron energy profile in N-I-N diode	70
3.20	3D electron velocity distribution in N-I-N diode	71
3.21	1D averaged electron velocity profile in N-I-N diode	71
4.1	Definition of the impact parameter	77
4.2	Comparison of screened and unscreened Coulomb potential	80
4.3	Comparison of BH and CW scattering potential	83
4.4	Propagation using Velocity Verlet algorithm	91
4.5	Velocity Verlet propagation with random scattering events	92
4.6	Particle energy under reflection off surface	94
4.7	Total energy during propagation in an elliptical orbit	96
4.8	Rutherford scattering trajectories	97
4.9	Comparison of simulated scattering angle dependence of impact parameter with Rutherford result	98
4.10	Simulated conservation of kinetic energy as function of impact parameter	100
4.11	Comparison of the mesh interpolated and analytical fields	106
4.12	Short range field models	107

4.13	Scattering angle dependence upon impact parameter using short range field models	108
4.14	Scattering angle vs. impact parameter for modified potential with variety of cut-off radii	109
4.15	Simulated consercation of energy for short-range field models	110
4.16	conservation of kinetic energy over collisions between positive and negative charges using modified potential with a range of cut-off radii	112
4.17	Nearest neighbour selection for force evaluation	113
4.18	Schematic 3D atomistic simulation structure for recovery of bulk mobility	115
4.19	3D potential in resistor structures	117
4.20	1D averaged potential profiles	118
4.21	Donor and electron concentration profiles	119
4.22	Electric field profiles	120
4.23	Average electron velocity profiles	120
4.24	Reproduction of low field concentration dependent bulk electron mobility in silicon at 300K	122
4.25	Kinetic energy during ionized impurity interaction	123
4.26	Electron kinetic energy distribution in space	124
4.27	Phonon emission during ionized impurity interaction	125
4.28	3D potential profile throughout 'atomistic' N-I-N diode	127
4.29	Agreement between large 'atomistic' and 'continuous' N-I-N devices . .	128
4.30	Comparison of 'atomistic' electron concentration with continuous simulation	129
4.31	Comparison of 'atomistic' and continuous electron energy profiles	130
4.32	Comparison of 'atomistic' and continuous electron velocity profiles . . .	131
4.33	Random donors highlighted as effecting energy and velocity profiles . . .	131

5.1	Node centred continuous charge assignment	136
5.2	Comparison of continuous charge and CIC point charge assignment . . .	137
5.3	Discrepancy combining continuous and discrete doping	139
5.4	Assignment of continuous charge via CIC weighting	140
5.5	Unique regions over which integration around a node is performed	141
5.6	Continuous charge assignment in node aligned mesh	143
5.7	Discrepancy combining continuous and discrete doping	144
5.8	Assignment of continuous charge via CIC weighting	145
5.9	Donor profiles with and without CCIC charge assignment	146
5.10	Donor profiles with and without CCIC charge assignment on a more coarse mesh	147
5.11	30nm MOSFET with trapped charge	149
5.12	Caughey-Thomas mobility model fit to the bulk MC result	151
5.13	Fit of the velocity-field curve to bulk MC simulation	153
5.14	Comparison of DD and self-consistent MC	154
5.15	Compare electron concentration profile between DD and self consistent MC	154
5.16	Compare current from DD and frozen field MC	155
5.17	Percentage current reduction in 30nm MOSFET	156
5.18	Electron current density with and without trapped charge, $V_G = 0.4V$. . .	159
5.19	Electron current density with and without trapped charge, $V_G = 1.0V$. . .	160
5.20	Electron current density with trapped charge from DD simulation	161
5.21	Electron current density profiles	162
5.22	Electron density profiles from DD and MC	164
5.23	Percentage current reduction in 20nm MOSFET	165
5.24	Percentage current reduction in 10nm MOSFET	166

5.25	Percentage current reduction associated with position dependent trapped charge	167
5.26	Current variation in 50 'atomistic' MOSFETs	171
5.27	Correlation between DD amd MC drain current	173
5.28	Percentage current variation in 50 'atomistic' MOSFETs	174
5.29	Percentage variation from DD vs MC	175
5.30	DD Current density throughout device 30	177
5.31	MC Current density throughout device 30	177
5.32	DD Current density throughout device 29	179
5.33	MC Current density throughout device 29	179
5.34	DD Current density throughout device 38	181
5.35	MC Current density throughout device 38	181
5.36	Current density throughout device 41 from frozen-field MC simulation . .	184
5.37	Current density throughout device 41 from self-consistent MC smiulation	184
5.38	Potential profile at interface from DD simulation of device 41	185
5.39	Potential profile at interface from self-consistent MC simulation of device 41	185

List of Tables

2.1	ITRS Roadmap scaling trends	11
3.2	Optical phonon energies	44
3.4	Experimental drift mobilities in silicon	52
3.6	Band independent parameters and calibrated values for the minimum lying X -valley	73
3.8	Parameters for the L and Γ valleys	74
4.1	Dimensions and doping concentrations of resistor devices for recovery of bulk mobility	116
4.3	Average field and velocity throughout bulk region	121
5.1	Unique regions and limits of integration associated with cells contributing to the assigned charge at a node	142
5.2	Unique regions and limits of integration associated with cells contributing to the assigned charge at a node	145
5.3	MOSFET device dimensions	150
5.4	Devices 29, 30 and 38 have similar values for I_D from DD simulation, but those from MC simulation show large variation.	176

Chapter 1

Introduction

Following the forecast of the International Technology Roadmap for Semiconductors (ITRS) [1], silicon devices with channel lengths of 7 nm are expected to be in mass production by the year 2018. Currently, conventional Si MOSFETs in the 90 nm technology node with channel lengths around 50 nm are commercially available in home PCs [2], while the demonstration of research devices corresponding to the next generation technology nodes has been continually reported [3–11]. Fabricated 30 nm physical gate length conventional silicon MOSFETs were reported in 2000 [3] showing good performance in terms of off current, subthreshold slope and drive current. Experimental results of a high performance 35 nm physical gate length conventional silicon MOSFET were later presented in 2001 [4] and showed improved drive current compared to [3]. A 20 nm physical gate length MOSFET, which is a scaled version of the 30 nm device reported in [3], was published in [5] showing “healthy IV characteristics”. Reports on 16 nm , 15 nm and 14 nm physical gate length conventional silicon MOSFETs, [6–8] quickly followed. The successful fabrication and characterisation of the smallest conventional silicon MOSFET with a 4 nm physical gate length [9] was reported in 2003. These ultra small devices show proper operation although some of them need further optimisation in order to meet the

requirements of the the 2004 edition of the ITRS [1]. The 2004 edition of the roadmap, however, predicts that conventional MOSFETs will be replaced after the 65 nm technology node with MOSFET architectures that have improved drive current and electrostatic integrity, such as thin body SOI and double gate transistors. Such novel devices are likely to remain in production until the end of the roadmap, in some cases in competition with conventional MOSFETs. These novel device structures with ultra small gate lengths have been successfully demonstrated with the publication of measured results from a 10 nm silicon finFET [10] and a 6 nm gate length ultra-thin silicon channel SOI device [11]. The successful demonstration of the above devices verify that aggressively scaled silicon devices can be fabricated, and with refined fabrication process may well meet the requirements of the ITRS.

However, the fabrication process of all semiconductor devices inherently introduces random deviations in the ideal device structure from one device to another. Even small atomic scale variations can significantly effect the characteristics of ultra small devices, leading to a large statistical distribution in device parameters over a fabricated ensemble. The corresponding large mismatch in device parameters adversely effects circuit operation [12, 13]. This poses a significant limitation on continued device scaling and future gigascale integration (10^9 or more devices on a single chip) [14]. This means that, although the successful demonstration of one device may be attainable, the mass production of such devices with a small enough statistical parameter distribution in order to allow their successful integration may not be possible. It is therefore necessary to be able to accurately predict the statistical distribution of parameters associated with random variations in future ultra small devices in order to assess the impact upon the next generation of circuits and systems [15–17].

Within conventional silicon MOSFETs, the three most important sources of intrinsic parameter variations are Oxide Thickness Fluctuations (OTF), Line Edge Roughness

(LER) and Random Dopants (RD). OTF refers to the random variation in the gate insulator thickness incurred during its growth or deposition process. The resulting rough interface between the semiconductor and the insulator locally effects the gate capacitance and alters the surface potential. This in turn alters the inversion layer charge for a given gate bias and leads to variation in threshold voltage V_t . Variation in the semiconductor/insulator interface will also induce local variations in surface roughness scattering through the variations in electrostatic potential which is expected to add to the device parameter variation. LER is caused by the random variations in gate edge lithography due to the molecular non uniformity of the resist material and imparts a spatially varying pn junction at either end of the channel [18,19]. Variation in the position of the edge of these junctions along the channel width leads to a variable channel length and can significantly effect V_t , I_{on} and I_{off} more strongly than OTF in most cases [20,21]. However, the most influential factor introducing intrinsic parameter fluctuations within the current technology node is the variation in electrostatic potential associated with the random number and position of discrete impurity atoms. The simulation study of the random dopant induced intrinsic parameter fluctuations is the focus of this thesis.

1.1 Random Dopants

The ultra small MOSFETs that have been reported [3–9] require extremely high doping concentrations in the channel in order to attain good electrostatic integrity. Following the scaling rules set out in the ITRS [1], devices with channel lengths less than $10nm$ require channel doping in the order of $10^{19}cm^{-3}$ and above. Even at such large doping concentrations, the small volume of the channel in future minimum-geometry transistors results in only tens of dopant atoms within the active region. The statistical nature of the ion implantation and diffusion leads to variation in both the number and position

of the dopant atoms within the channel of the individual transistors. The variation in numbers closely follows a Poisson distribution with the number of dopants from device to device fluctuating about the expectation number N with a standard deviation of \sqrt{N} . Variation in the number of dopants within the channel alters the inversion charge density from device to device at a given applied gate bias and consequently alters V_t . Threshold voltage variation due to fluctuations in doping charge density was first recognised and considered as a potentially limiting factor in device scaling in 1972 [22, 23].

Additionally, the random position of each dopant atom results in a unique charge configuration. The loss or inclusion of a single dopant atom within the channel can significantly alter the electrostatics and the characteristics of a devices with so few dopants. Within the channel, the potential energy landscape associated with electrons in an ultra small n -channel MOSFET is dominated by local peaks corresponding to the acceptors and lower 'valleys' in-between. Electrons readily flow through the valleys where the electrostatic barrier is lower, forming current percolation paths. The availability of such paths depends largely upon the 3D dopant configuration.

Predicting the magnitude of random dopant induced device parameter variation in ultra small devices is essential in order to assess the ability to reliably integrate such devices in circuits. Simulation of statistical samples of microscopically different devices provides a means of predicting the distribution of device characteristics. Simulation also offers an effective means by which individual sources of intrinsic parameter fluctuations can be studied independently in order to assess their relative importance. In studying random dopant induced effects, simulation of semiconductor devices in which the individual dopant atoms are resolved becomes a requisite tool. Extensive work on simulating statistical sets of randomly configured devices in order to characterise the statistical variation in device characteristics has been mainly performed using 3D Drift Diffusion (DD) simulations [20, 21, 24–26]. Such simulations are limited by the validity of the local transport

approximation inherent to the DD approach and can only capture the impact of a unique arrangement of dopants through the electrostatic influence upon the inversion layer carrier density. It is expected that variations in number and position of random dopants, acting as Coulomb scattering centres, will readily effect transport throughout the channel as a result of variations in position dependent scattering. This effect is unobservable in DD simulations and, as a result, much of the published results on random dopant induced device parameter variation only reveal part of the full inherent variation. In order to assess the full scale of device parameter variation, simulation techniques must also properly include a fundamental description of charge transport that can capture the position dependent scattering associated with a random arrangement of charges in the channel. The Monte Carlo (MC) method is such a technique that, with a careful modification to the way the Coulomb scattering is treated, can be used to investigate the full impact of random dopants on device parameter variations, including both electrostatic and transport effects.

1.2 This Work

Within this work, the simulation of transport variations associated with the random configuration of discrete ionized impurity atoms, and the corresponding increase in the simulated intrinsic parameter variations in nano-scale conventional silicon *n*-MOSFETs, are studied in detail. An *ab initio* Ensemble Monte Carlo (EMC) simulation tool is developed that can accurately account for the variations in the ionized impurity scattering of electrons through their real-space trajectories in the electric field associated with an arbitrary arrangement of dopant atoms. Care is taken to ensure an accurate description of electron-impurity scattering and also electron-electron scattering via the same real-space trajectory approach. Methodology has been developed which allows such an *ab initio*

approach to scattering to be included within traditional EMC simulation. The developed simulation tool has been used to study the transport contribution to random dopant induced device parameter variations in nano scale MOSFETs. Through careful comparison with DD simulation, the electrostatic influence of random discrete charges is separated from the additional effect of transport variation, and the relative importance of the two effects in different operating regimes is highlighted. It is shown that the inclusion of transport variation does indeed increase the random dopant induced device parameter variation compared to DD simulation.

This work is laid out as follows:

In chapter 2, the random dopant effects are discussed and the related experimental and simulation research on random dopant induced intrinsic parameter fluctuations is reviewed. A brief summary of the other major sources of intrinsic parameter fluctuation that effect device performance is also introduced. These additional sources of fluctuation may readily be incorporated within the developed MC simulation tool to further this work.

The standard Monte Carlo method is introduced in chapter 3, where details of the models adopted within this work are described. The calibration to bulk Si transport properties and the validation of the models in the low-field regime used in this work is reported. The description of the 3D device simulator is also given with further verification through the simulation of an example N-I-N diode.

In chapter 4, the 3D device simulator is extended to include an *ab initio* description of Coulomb scattering from dopants and carriers. The implementation of this approach is discussed, paying particular attention to the accurate integration scheme for the equations of motion of electrons. The validity of the developed approach is shown through the reproduction of the experimentally measured concentration dependence of the low field electron mobility in bulk silicon.

Having developed a suitable means of incorporating *ab initio* scattering from random

discrete charges, in chapter 5 the results of simulations designed to investigate the mobility variations associated with random arrangements of ionized dopants within MOSFETs are presented. Initially, investigation of the impact of a single trapped charge on the transport is highlighted in comparison to DD simulation. Next the on current variation in a set of randomly doped devices is considered. The thesis conclusions are drawn in chapter 6 where suggestions for future work are also made.

Chapter 2

Intrinsic Fluctuations and Device Parameter Variations

2.1 Introduction

In the past, the main source of parameter fluctuations in CMOS devices were due to variations in the process parameters leading to slowly varying differences in doping profiles, layer thickness and device dimensions from wafer to wafer or across a single wafer. Tighter control of the processing conditions were able to reduce such “extrinsic” parameter fluctuations. With the reduction of the CMOS devices to decanometer and nanometre dimensions in the present and future generation technology nodes, “intrinsic” parameter variations due to the discreteness of matter and charge becomes increasingly important and start to dominate the device parameter variation. Intrinsic parameter variations become especially problematic in the design of ultra small devices where microscopic differences in the structure of otherwise macroscopically identical devices lead to an increasing statistical distribution in device parameters, such threshold voltage V_T and off current I_{off} , over a fabricated ensemble. The spread in device parameters, or mis-

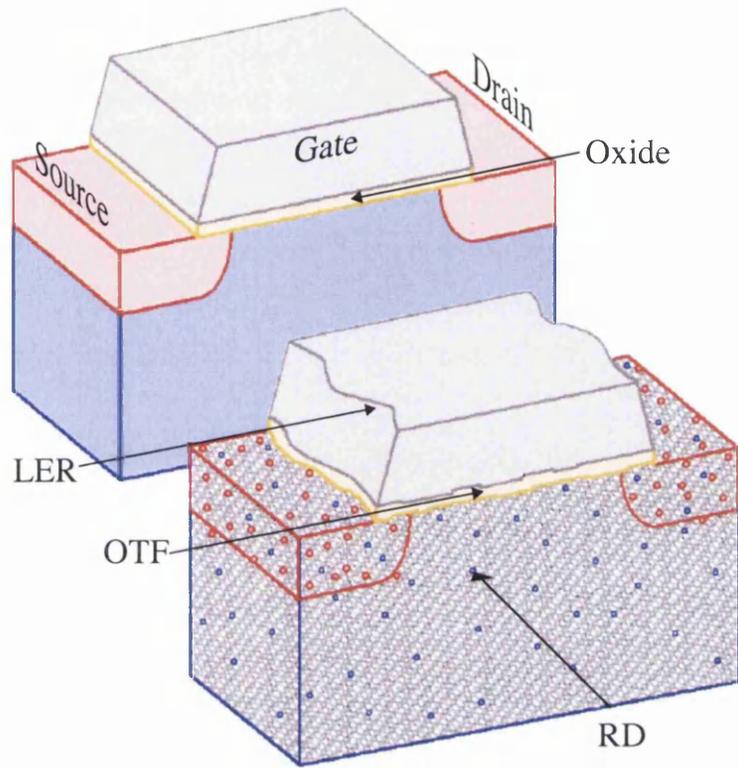


Figure 2.1: (Top) Ideal device description with uniform charge distribution and well defined edges. (Bottom) Realistic ultra small device with microscopic variation in discrete charges and lithography.

match between devices, is detrimental to the successful integration of such devices in both analogue and digital circuits and represents a limit to future scaling trends. Developing device designs and architectures that are resistant to intrinsic parameter variation is then a primary concern. It is therefore important to be able to accurately predict the magnitude of the intrinsic device parameter variation for different device designs in order to help limit device mismatch and achieve optimal circuit performance and reliability in the future generation technology nodes.

Figure 2.1 illustrates the differences between an ideal macroscopic device typically considered in practical device simulation and the corresponding structure of a real microscopically discrete device. In the ideal case, the doping and the corresponding charge

density is represented by a continuum distribution. However, and as will be discussed in more detail, in ultra small devices the number of individual ionized impurity atoms is small and the effect of variation in their number and position is observable from device to device. The lithography defining the gate edges has inherent line edge roughness (LER) which becomes comparable to the device dimensions and introduces random variation in the transistor geometry. Finally, the interface roughness on the scale of one or two atomic layers introduces random variations in the gate dielectric thickness.

Therefore the simulation of ideal devices with straight boundaries and continuous doping is simplistic and will not reveal anything of the extent of the intrinsic parameter variation. Simulation of an ensemble of microscopically different devices in terms of discrete doping distribution, interface and LER patterns must be carried out in order to statistically characterise the corresponding intrinsic parameter fluctuations.

The magnitude of the effects associated with discrete random dopants (RD), line edge roughness (LER) and oxide thickness fluctuations (OTF) depends upon the device structure. For this work, conventional silicon *n*-MOSFETs are considered as they are the main focus of the semiconductor industry's scaling targets at present and in the near future. In section 2.2, random dopant induced device parameter variations in MOSFETs are introduced and the estimation of the intrinsic parameter variation using Drift Diffusion (DD) simulation is discussed. Both the limitations of the DD approach to account for transport variations associated with random dopants and the role of Monte Carlo (MC) simulation to properly incorporate this important effect are discussed. Subsequent to this, LER and OTF as sources of intrinsic parameter fluctuations are discussed in section 2.3 before additional sources of fluctuation that will become important in future devices are considered in section 2.4.

Year of Production	2005	2007	2010	2013	2016	2018
Printed Gate Length (<i>nm</i>)	45	35	25	18	13	10
Physical Gate Length (<i>nm</i>)	32	25	18	13	9	7
Equiv. Oxide Thickness (<i>nm</i>)	1.1	0.9	0.7	0.6	0.5	0.5
Max Doping $V_T = 0.4V$ (cm^{-3})	2.5×10^{18}	5.0×10^{18}	9.0×10^{18}	1.8×10^{19}	3.0×10^{19}	N/A
Approx Number of Dopants	56	50	35	27	17	N/A
Random variation in dopants (%)	13	14	17	19	24	N/A

Table 2.1: MOSFET dimensions taken from the ITRS 2004 update edition [1]. Doping concentrations until 2016 are taken from 2002 update edition [27] before the emphasis on SOI devices. Expected number of dopants within the channel depletion region considering a square channel is also tabulated.

2.2 Random Dopants

Currently, the 2004 edition of the International Technology Roadmap for Semiconductors (ITRS) [1] sets the scaling trends for semiconductor devices until the year 2018. Table 2.1 lists the target ITRS channel lengths for the next generation MOSFETs until the end of the roadmap. The corresponding large channel doping concentrations required to scale conventional MOSFETs to such dimensions are also tabulated. Despite the high doping concentrations, there are only expected to be tens of dopants within the active region of the transistors. Due to the random nature of the ion implantation and diffusion process, the limited number of dopants leads to significant statistical variation in their number, as well as position, between different devices.

The limitation on device scaling imposed by the random discrete dopant distributions was proposed in the early seventies [22, 23, 28]. From the basic MOSFET theory, the threshold voltage V_T of a long channel MOSFET can be written as

$$V_T = V_{fb} + 2\psi_B + \frac{\sqrt{4\epsilon_{Si}qN_a\psi_B}}{C_{ox}} \quad (2.1)$$

where V_{fb} is the flat-band voltage, ψ_B the difference between the Fermi level and intrinsic

Fermi level, ϵ_{Si} the permittivity of silicon, C_{ox} the oxide capacitance and N_a the nominal substrate doping concentration. The square root term equals the total depletion charge, variation in which leads in first order to variation in V_T .

When considering purely random positions for the dopant atoms, the statistical distribution of the number of dopant atoms from device to device is assumed to follow a Poisson distribution since the location of individual impurity atoms are considered independent [23, 29]. Therefore, for an expected number of dopants \bar{N} within any volume, the actual number will statistically fluctuate about the mean with a standard deviation of $\sqrt{\bar{N}}$. For ultra small devices with few dopant atoms, this represents a significant percentage variation in the number of dopants and the depletion charge. This is also tabulated in table 2.1 where variations on the order of 10 to 25 percent are seen. In turn this results in significant variation in V_T . Excessive variation in V_T and drive current I_{on} limits the operation of analogue circuits while in ultra small devices, the magnitude of the device variation can be large enough to lead to malfunctions in digital electronics [30–33]. Such mismatch between devices is therefore crucial to predict in order to be able to foresee its impact on the device, circuit and system design.

2.2.1 Experimental Observations

Experimental observation of device variation and mismatch has been reported [29, 34–38], prompting studies of the variation in V_T and I_{on} on the subsequent degradation in circuit performance. At the time, this was addressed using mainly analytical models matched to the experimentally observed variations [34]. Analytical models assumed forms for the statistical variation of multiple mechanisms, such as oxide thickness, oxide trapped charge and line edge roughness. The statistical influence of random dopants were explicitly included as a source of V_T variation in [29], while in [34] a more general variation in

channel mobility was considered.

The direct experimental evidence of threshold voltage fluctuations associated with variation in random dopant numbers was first reported in 1993 [39–41]. The authors have studied a sample of over 8,000 MOSFET devices on a single chip, showing that the V_T variation was independent of the location of the device on the chip and therefore occurred randomly within devices. Moreover they have demonstrated that the distribution in V_T follows a Gaussian distribution with a well defined standard deviation in threshold voltage σV_T . The distribution in experimentally observed dopant numbers between devices was also observed to follow a Gaussian distribution. Line Edge Roughness and Oxide Thickness Fluctuations were considered minimal in the fabricated devices and ignored, and it was concluded that the random variation in dopant numbers were primarily responsible for the variation in V_T . Subsequent to this, careful measurement of the devices at differing drain and substrate voltages revealed that the dopant configuration within the channel volume, as well as the actual number of dopants, were important factors in altering V_T [42].

2.2.2 Modelling Approaches

Given the importance of random dopant induced intrinsic parameter variations, models to predict the magnitude of the standard deviation in threshold voltage σV_T have been developed. One simple approach [23, 28] includes the subdivision of the channel into an array of square MOS capacitors with length equal to the gate depletion width at the threshold voltage of the device. Variation in the total depletion charge within each MOS capacitor was assumed to follow a Poisson distribution and was considered to be uniformly distributed throughout. The probability that a series of such capacitors were inverted at a given gate bias to form a conducting path from source to drain was determined. This

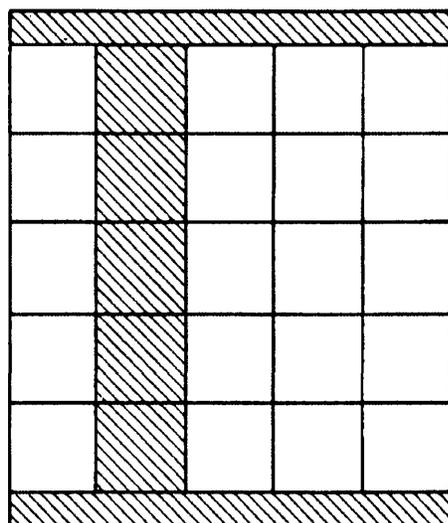


Figure 2.2: Simple model to account for the variation in substrate doping density within the channel of a MOSFET [23]. A conducting path from source to drain (top to bottom) formed by the alignment of a series of inverted MOS capacitors is highlighted.

is illustrated in figure 2.2 where the most likely conducting path is shaded. This very basic model was the first attempt to quantify the variation in threshold voltage and since then more complex analytical models have been developed [29, 36, 43–46]. Among the additional factors taken into account in these models is the variations in dopant distribution as a function of depth within the channel which was initially neglected but plays an important role [46].

Numerical device simulations are limited by the accuracy and the validity of the simulation approach but provide in general a far more accurate means of determining the magnitude of the random dopant induced intrinsic parameter variation through the simulation of an ensemble of microscopically different devices. They automatically capture complex effects associated with the real space variations. Although the numerical device simulation represents the most accurate way to investigate intrinsic variations in novel devices, it can be computationally very exhaustive. Despite the three dimensional nature of the

effects associated with random discrete dopants, initially two dimensional drift-diffusion (DD) simulations were used to study random dopant induced MOSFET parameter fluctuations [46–48]. Variation in the doping concentration was randomly assigned to the simulation mesh following a Poisson distribution of dopants with a mean concentration equal to the continuous doping concentration in the device. Through such simulations it became clear that the random dopant induced variation in the doping density results not only in V_T variations, but also in mean values of the threshold voltage different from the values obtained from the continuous doping simulations. This was attributed to the variation in surface potential resulting in early percolation of current through the channel. It was also found that the variation in doping close to the interface dominates the intrinsic parameter variations [47]. Physically, the random number and position of dopants alter the potential distribution throughout the channel, locally altering the barrier height. Valleys within the potential energy profile allow parts of the channel to turn on earlier than others, forming current percolation paths and lowering the average value of V_T . While 2D and 3D simulations show comparable results for σV_T , with 3D simulations giving a slightly smaller standard deviations [47], the threshold voltage shift could only be accurately modelled in 3D due to the 3D nature of the current percolation paths [46].

Following this, 3D 'atomistic' simulations incorporating individual random discrete dopants were performed [49]. In contrast to merely varying the charge density at the discretization grid nodes, individual dopant atoms were incorporated and the corresponding charge density assigned to the discretization mesh. The number of dopants was selected from a Poisson distribution, with mean equal to the expected number of dopants as given by the integral of the doping concentration over the doping volume, and positioned at random. This was extended via a rejection technique to accurately position dopant atoms at silicon lattice sites giving, on average, the required doping concentration [50]. Recently, dopant distributions coming from 3D atomistic process simulation [51,52] have been used

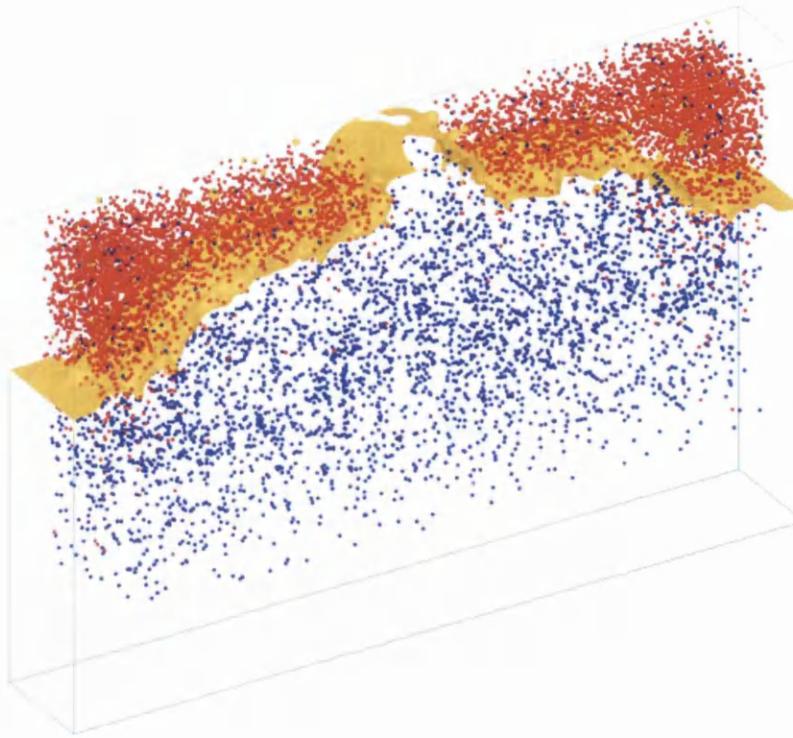


Figure 2.3: Atomistic dopant distribution resulting from 3D process simulation and electron equi-concentration contour from subsequent DD simulation [53].

to import the exact dopant positions for subsequent atomistic device simulations [53–55]. This represents the most accurate method of implementing the statistical fluctuations in dopant number and position prior to device simulation. Figure 2.3 shows the distribution of individual donor and acceptor positions, resulting from discrete dopant Kinetic Monte Carlo process simulation, and an electron equi-concentration contour from subsequent DD simulation [53].

Three dimensional simulations using the DD approach have since become the most common and efficient method to study random dopant effects and have been applied to a variety of device structures [24, 25, 50, 56, 57]. Within this model, problems arise when resolving the individual discrete dopant potential on fine discretization meshes associated with ultra small devices. The resolution of the singular Coulomb potential associated with

a point charge of an ionized impurity atom leads to unphysical charge carrier localisation within the DD solution. The implementation of correction schemes in order to avoid this artificial localisation effect has recently been reported and has allowed drift diffusion simulations of ultra small devices to be carried out [54, 58–60]. In addition, quantum mechanical effects are often included through the density gradient (DG) formalism [61] within DD simulations to account for the carrier confinement in the inversion layer when simulating small devices [62, 63].

2.2.3 Single Charge Trapping

Related to the discrete random dopant induced parameter variation is the variation in the MOSFET current induced by the trapping/de-trapping of single charge carriers in defect states within the channel. For example, trapped electrons in n -channel MOSFETs alter the potential landscape within the channel resulting in a local reduction of the charge carrier density and a corresponding reduction in the current [64]. Unlike random dopants, this is a dynamic process, albeit on timescales much longer than the switching time of devices, that results in the random telegraph signal (RTS) in the device current. The magnitude of the current reduction associated with a single trapped charge has been studied in a similar manner to random dopants using three dimensional DD simulation [65, 66].

Combined with random dopant simulations, where current conduction through percolation paths dominates at low gate voltages, the effect of a single trapped charge can be significant [25]. Trapping of a single charge within a percolation path presents a barrier to current flow and can cause significant RTS noise. Experimentally observed RTS amplitudes of greater than 60% have been reported in decanometer devices [67] while simulation studies that fail to properly account for the 3D discrete nature of the doping charge underestimate the magnitude of the variation [68]. Accurate simulation of discrete

dopants and trapped charges is therefore important in order to estimate the magnitude of noise in device operation as a source of dynamic device variation.

2.2.4 Limitation of Drift Diffusion Simulation

As stated already, much of the statistical work on random dopant induced device parameter variation has been carried out using the drift-diffusion approximation. This is a local transport model which can self consistently solve for the carrier concentration and current continuity for a given background charge distribution and externally applied biases [69]. In doing so, values for the macroscopic diffusion constant and mobility are used which correspond to steady-state bulk values. Empirical mobility models fitting experimentally observed concentration dependent and field dependent mobility can also be used to increase the validity of the model [70]. However, the mobility values used within the solution are associated only with the local field and doping concentration and as a result DD simulation is incapable of capturing non-equilibrium transport [71]. The model holds where fields are slowly varying such that equilibrium conditions are well approximated. However, the field associated with a discrete impurity is rapidly changing and does not fit this description. Similarly, the local variation in doping concentration associated with the resolution of discrete dopants on a finely discretized mesh can introduce unphysical local variation in the mobility. These make the direct use of concentration and field dependent mobilities in atomistic DD simulations unsuitable. In order to circumvent the above problems, the field and doping concentration dependence of the mobility obtained from continuous doping simulation is used directly in the atomistic simulation at equivalent bias conditions [24]. In this manner, the discrete charges effect the DD simulation through the local increase or reduction in carrier concentration associated with local variations in electrostatic potential due to the discrete charges. This in turn effects the

current.

In addition to the electrostatic modulation of the carrier concentration, the discrete charges act as Coulomb scattering centres. Variation in charge transport due to variation in Coulomb scattering from a non self averaging array of discrete impurities is expected to have a significant contribution to the current reduction in transistors with only a few dopants. Similarly, variation in transport associated with a trapped charge is expected to contribute to the reduction in current flow, however it is still unclear as to whether the mobility or the local carrier density modulation is of primary importance [72–74]. The mobility models employed within DD cannot capture transport variations due to localised Coulomb scattering centres. As a result, the magnitude of the intrinsic parameter variation is underestimated by DD simulations since the intrinsic transport variations are neglected. The magnitude of the effect of transport variations between devices due to varying Coulomb scattering is thus an important question which this work addresses.

2.2.5 Monte Carlo Simulation

Including transport variations requires a non-equilibrium transport model that can account for scattering from an arbitrary arrangement of scattering centres in three dimensions. The Monte Carlo device simulation approach is suitable for studying transport in the presence of random discrete dopants provided that modifications are made to treat Coulomb scattering from discrete charges through the real space trajectories of the carriers. Previously, electron-electron scattering [75] and electron-ion scattering [76] have been treated for an arbitrary arrangement of scattering centres through the inclusion of a molecular dynamics force evaluation within the traditional Monte Carlo simulation. However these simulations were of bulk material and relied upon periodic boundary conditions imposed on the simulation domain. Methods have been developed to incorporate the molecular dynam-

ics approach in Monte Carlo device simulation in order to study random dopant induced variation and electron-electron interactions in MOSFETs [77–81]. Some statistical study of simulations of between 15 and 30 randomly configured devices using this *ab initio* method again concluded that variation in dopant position, as well as number, is an important factor affecting the device parameter variation, with impurities closer to the oxide interface playing a more dominant role [78,81]. The *ab initio* method of including ionized impurity scattering is itself left for discussion in chapter 4 and is of primary importance in this work.

2.2.6 Random Dopants in Non-Conventional Si Devices

Simulation of discrete random dopant effects has so far been focused on parameter variations in conventional MOSFETs. In the most recent editions of the ITRS [1], it has been postulated that the conventional MOSFET device will be replaced by fully depleted thin body SOI MOSFETs somewhere beyond 2007 and later on with multiple gate device architectures. Such devices have little or no channel doping and, as a result, unwanted discrete impurities or trapped charges within the channel may significantly affect the device operation [66]. The random position of the highly concentrated source and drain dopants, however, will alter the effective position of the *pn* junction leading to significant effective channel length variation in ultra short channel devices [82]. Therefore random dopant induced parameter variation, being one of the most important limiting factors in current devices, will remain relevant until the end of the ITRS. The accurate modelling of random dopant induced variation therefore remains very important until the end of the roadmap and beyond.

2.3 Line Edge Roughness and Oxide Thickness Fluctuations

In addition to the random discrete dopant distribution which introduces variation in the device characteristics, the fabrication processes introduce variations in the gate dimensions and in the gate insulator thickness. Variation in the lithography defining the gate edge alters the physical channel length while variation in the gate insulator thickness directly effects the local gate capacitance and hence local surface charge density. These effects in turn contribute to the variation of the threshold voltage and the current-voltage characteristics in an ensemble of nominally identical devices. Both LER and OTF are the main sources of intrinsic parameter variation commonly studied in addition to random dopants.

2.3.1 Line Edge Roughness

For a particular generation of lithography, the resolution of the lithographic process is primarily limited by the mask and the resist material, and not the lithographic tools [83, 84]. Local variation in the density of the resist material through the formation of polymer aggregates was originally observed for a variety of different resist materials [85]. Aggregates form from the entanglement of the polymer chains and range in size between 10 – 30 *nm* in diameter depending upon material and light or electron beam exposure [83]. The aggregate size alters the response of the resist to the development process, with large aggregates dissolving less rapidly. This results in variation in the developed resist features that correlates with the distribution of such aggregates. This variation is then directly transferred to variation in the underlying material [83–85]. Experimentally observed variation in the surface of an exposed resist material and the subsequent roughness in a patterned line is

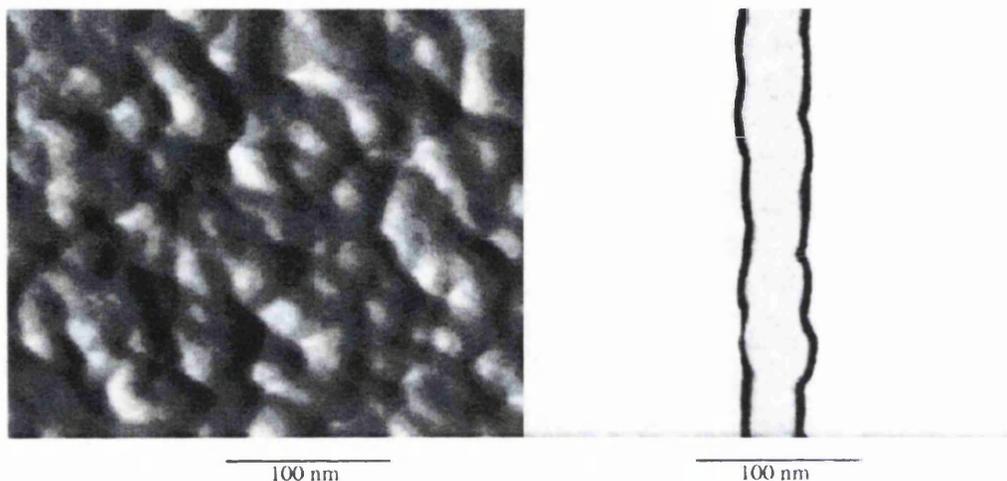


Figure 2.4: (left) The surface of an exposed resist showing the aggregates that lead to the loss in edge resolution of a pattern, and (right) the subsequent LER produced in the lithography of a line. Taken from [84]

shown in figure 2.4 [84].

Crucially within the MOSFET structure, the resolution of the gate sidewall profile and of the source and drain implantation regions are limited by LER. Variation in lithography is transferred to variation in the source/channel and drain/channel junctions following the ion implantation and diffusion processes. Independent LER variations in the source and drain profiles results in Line Width Roughness (LWR) [19] which corresponds to variation in the effective channel length. This effect will be significant in short channel devices where variation in channel length represents a significant percentage variation compared to the nominal channel length.

Much work has been done to statistically describe LER variation [18]. Typically LER is described by an RMS amplitude Δ and a correlation length Λ , with reported values ranging broadly between $3\Delta = 5 - 6nm$ and $\Lambda = 20 - 30nm$ [20]. The autocorrelation function of the experimentally observed rough line edge fits a Gaussian or Exponential autocorrelation function equally well. Noise in the LER measurements makes it hard

to distinguish which is a better fit, though the Gaussian form is more commonly used, resulting in less high frequency variation. This description has allowed the introduction of statistically accurate LER into statistical 3D DD device simulation [20].

Using continuous doping within the DD approximation, the effects of LER on V_T and I_{off} variation at low applied drain voltages are well represented, as in this regime the device is dominated by the device electrostatics. Monte Carlo simulation will better capture the non-equilibrium transport that occurs at high applied drain voltages where DD underestimates current. Simulations combining random dopants and LER (see figure 2.5) have been carried out to investigate the simultaneous effect of the two sources of fluctuation on the device parameter variation and show that, to a good degree, the two effects are statistically independent [20]. However, this treatment again avoids variation in scattering associated with the random discrete impurities. The effect of scattering from random dopants, particularly associated with low energy carriers entering the channel from the source, may be significant and again requires a generalised treatment of impurity scattering within the non-equilibrium Monte Carlo model.

2.3.2 Oxide Thickness Fluctuations

The formation of the gate, which currently involves oxidation and poly silicon deposition, typically introduces interface roughness on the order of 1-2 atomic layers at both the gate/insulator and semiconductor/insulator interfaces, figure 2.6 [86]. The corresponding oxide thickness fluctuations (OTF) alters the gate capacitance throughout the channel which in turn alters the surface potential profile from device to device. In future ultra small MOSFET devices, oxide thicknesses of less than 1 nm are required and variation by a single atomic layer introduces a significant percentage thickness variation. Like LER, OTF is statistically described by a Gaussian or Exponential autocorrelation function with

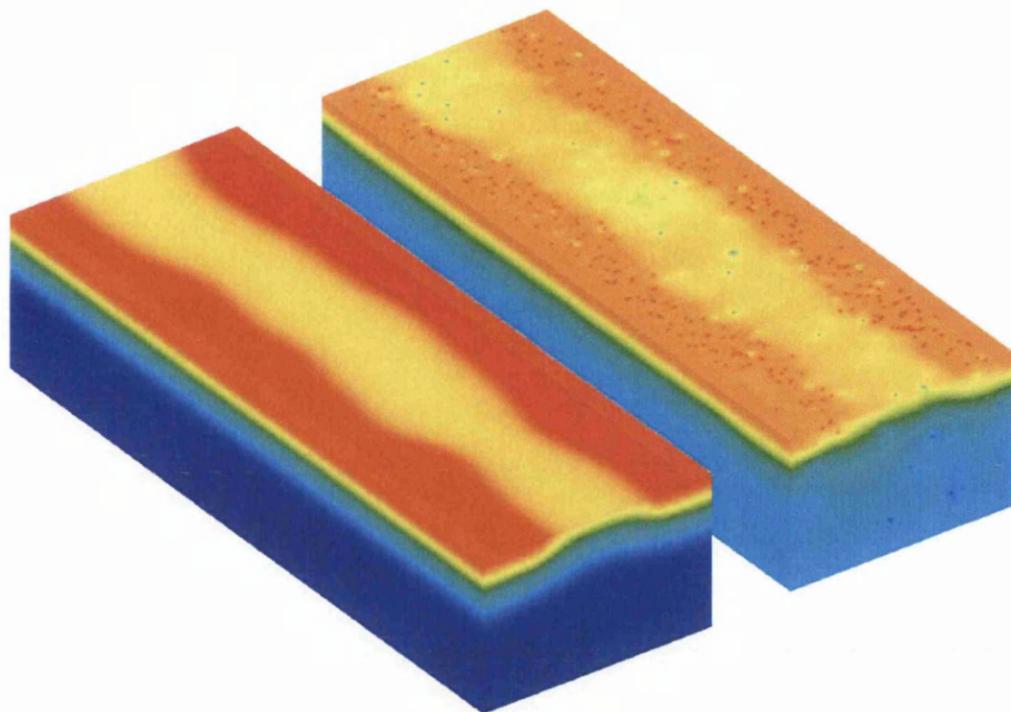


Figure 2.5: Electrostatic potential resulting from DD simulation incorporating LER (left) and both LER and random dopants (right). The same LER construction is used in each. Variation in channel width and local variations in potential associated with individual dopants are seen [20].

RMS amplitude Δ and correlation length Λ [87,88]. OTF has then been studied in a similar fashion to LER in 3D DD device simulation [89,90]. Simulations of OTF fluctuations including random dopants show that the two act as statistically independent sources of intrinsic parameter variation.

Within the framework of DD simulation, the varying oxide thickness electrostatically alters the surface potential and the local carrier density within the channel. Similar to the effects of random dopants, current percolation paths are formed where the potential variation due to the oxide thickness variations lowers the potential energy barrier to form a conducting path from source to drain. This results in variation in the threshold voltage as well as in a lowering of the mean threshold voltage similar to the description in section

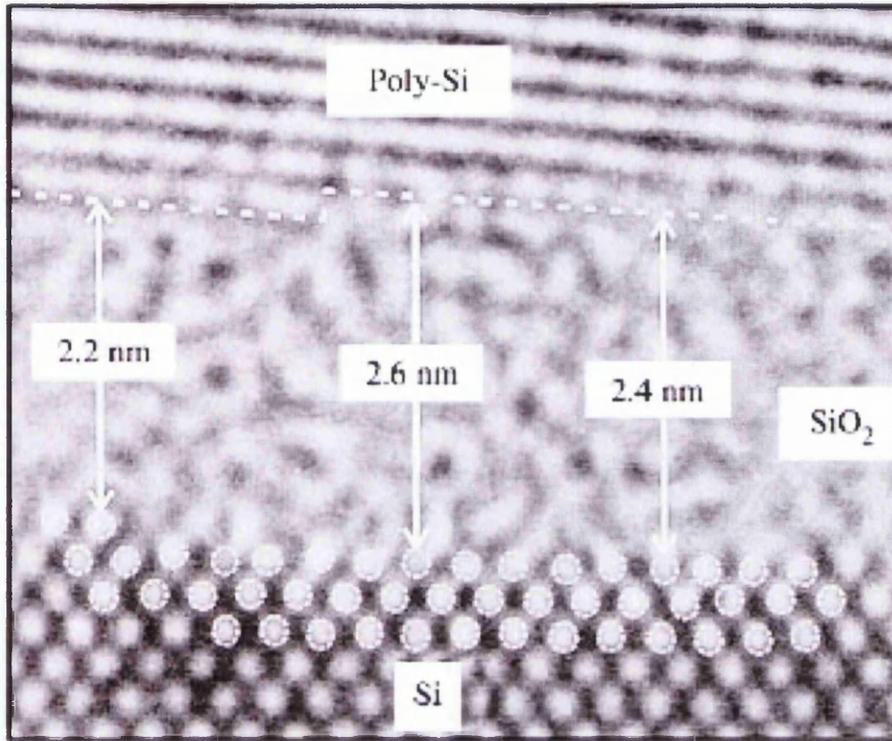


Figure 2.6: Image of gate/insulator/semiconductor interfaces. Variation in both interfaces is seen and results in a varying thickness of the intermediate oxide [86].

2.2.2.

However, local variation in the interface roughness pattern and the corresponding oxide thickness variation also effects variation in the surface roughness limited mobility which plays a crucial role in transport within the channel [87]. The varying electric field associated with the non-uniform surface potential acts to scatters carriers. Differences in the oxide interface pattern lead to variation in position dependent carrier scattering which is unaccounted for within the DD approach. It is expected that the variations in the interface roughness, in combinations with the additional variation due to scattering from surface potential fluctuations associated with the random oxide thickness pattern, will increase the intrinsic device parameter variation in the same way as Coulomb scatter-

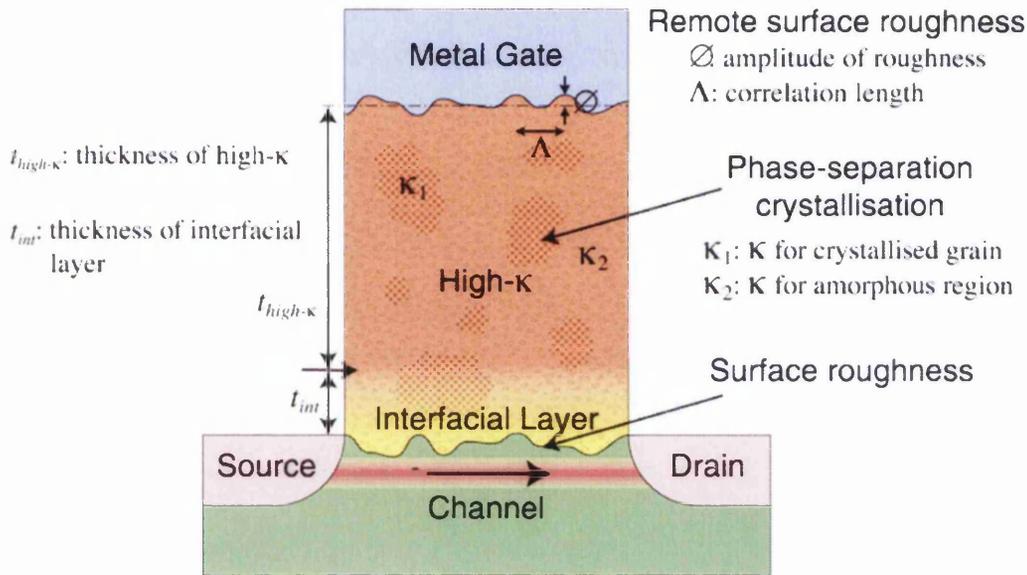


Figure 2.7: Illustration of the fluctuation problems faced when moving to high- κ materials in the gate stack. Surface roughness at the gate/insulator and insulator/semiconductor interface, as well as variation in the composition of the dielectric material, effect surface potential.

ing associated with random dopants. To include this in simulations requires an *ab initio* treatment of the interface roughness scattering similar to the *ab initio* random dopant scattering and can be suitable handled by 3D Monte Carlo simulation incorporating rough surfaces [91]. It is important within such simulations that the quantum mechanical confinement of carriers within the inversion layer is properly treated in order to self consistently capture the electron density and subsequent scattering from the varying potential.

2.4 Additional Sources of Fluctuation

For the ultra small devices at the end of the roadmap, oxide thicknesses of around 0.5 nm are required (see table 2.1). The reduction of the gate oxide thickness to increase the gate capacitance is limited by the requirements of minimising gate tunnelling current and by

the physical size of the silicon lattice. Moving towards gate insulator materials with a higher dielectric constant, high- κ materials, allows the gate capacitance to be maintained using much thicker insulators and at the same time controlling gate tunnelling. The high- κ materials may introduce additional sources of fluctuation that may become important sources of intrinsic parameter variation in ways that are only beginning to be addressed. An overview of the potential sources of random parameter variations associated with high- κ materials is shown in figure 2.7. Variation in the composition of the high- κ material associated with crystalline grains with a different dielectric constant compared with the amorphous high- κ material will further the surface potential fluctuations, increasing the parameter variations due to variations in charge density and transport. As before, a full treatment of these effects would require a proper 3D treatment of transport variations.

2.5 Conclusions

The most common sources of intrinsic device parameter variation, including discrete random dopants, line edge roughness and oxide thickness fluctuations, have been introduced. Accurate modelling of the magnitude of the effects associated with these fluctuation sources require full 3D simulations which have to date been carried out predominantly using the drift-diffusion approach. However, it has been noted that variations in transport due to the fluctuating potential landscape within randomly configured devices, associated with both random dopants and oxide thickness fluctuations, is expected to contribute to the intrinsic parameter variations. These effects are unobservable within the drift-diffusion approach and requires a proper treatment of non-local transport to reveal their impact on the statistical device parameter distributions. Further to this, additional sources of fluctuations inherent to high- κ materials, essential for future scaling of conventional MOSFETs, are expected to increase the surface potential variations, thus increasing transport associ-

ated variations.

Monte Carlo simulation is capable of capturing non-equilibrium transport and, including the impact of the 3D potential landscape on the real space carrier trajectories, should automatically account for such transport variations. Within the following chapters, the Monte Carlo method is discussed and a description of the incorporation of scattering from random discrete dopants is given. The developed simulator is then used to compare the results of device parameter variation including transport effects, with that of DD simulations which only capture the electrostatic influence of the different sources of intrinsic parameter variations.

Chapter 3

Monte Carlo Device Simulation

3.1 Introduction

Compared with drift-diffusion (DD), the Monte Carlo (MC) method [92–96] provides the most accurate means by which the classical influence of individual random discrete dopants on device operation can be examined. Such simulations should be three dimensional to be able to capture the impact of random numbers and, importantly, positions of dopants within the device. The MC method is a general mathematical tool for the numerical solution of a variety of problems not unique to semiconductor charge transport. In this chapter the application of the MC method to semiconductor charge transport is discussed, detailing the approximations used and their validity within the framework of the proposed work.

The classical description of charge transport in semiconductor materials is given by the solution of the semi-classical Boltzmann transport equation (BTE). The BTE is introduced in section 3.2. The solution of this equation could be achieved through the direct simulation of charge carrier trajectories, influenced deterministically by local electric fields and stochastically by random scattering processes. Throughout this work we

only consider electrons as majority charge carriers and consequently electron dynamics are briefly discussed through the semi-classical equations of motion. The relatively simple MC method applied in this work is described in section 3.3 in comparison with more advanced MC simulators. In section 3.4, the calibration and verification of the model through simulation of bulk silicon properties and comparison with both experimental data and more rigorous MC solutions is discussed. The important steps in moving from a simple bulk solution to a 3D self-consistent device simulation are discussed in section 3.5, followed by an example simulation of a 3D test structure.

3.2 Boltzmann Transport Equation

3.2.1 Single Particle Distribution Function

The Boltzmann transport equation (BTE) describes the evolution in time of the single particle distribution function, $g(\mathbf{r}, \mathbf{v}, t)$, which gives the probability of finding a particle with velocity between \mathbf{v} and $\mathbf{v} + d\mathbf{v}$ in the volume \mathbf{r} to $\mathbf{r} + d\mathbf{r}$ at time t . If known, the distribution function allows averages over a volume in the position-velocity phase space to be calculated as

$$\langle A(t) \rangle = \int \int A(\mathbf{r}, \mathbf{v}, t) g(\mathbf{r}, \mathbf{v}, t) d\mathbf{r} d\mathbf{v} \quad (3.1)$$

which are interpreted as an ensemble average. Knowing the distribution function thus enables the determination of, among other things, charge and current densities which are of primary importance in device simulation.

3.2.2 Classical BTE

The BTE is given by

$$\frac{\partial g(\mathbf{r}, \mathbf{v}, t)}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} g + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} g = \left(\frac{\partial g}{\partial t} \right)_{Coll} \quad (3.2)$$

We state it here, but a complete derivation may be found in [97]. The equation relates the time rate of change of the distribution function to both the number of particles with velocity \mathbf{v} flowing out of volume element $d\mathbf{r}$ at position \mathbf{r} , and the number of particles with acceleration $\frac{\mathbf{F}}{m}$ flowing out of a similar volume $d\mathbf{v}$ in velocity space. The terms on the left hand side comprise the drift terms, arising from the initial conditions and the motion in the electric field \mathbf{F} . The term on the right hand side is the collisional term and constitutes all possible mechanisms, other than the electric field, that alter the velocity. Equation 3.2 is a classical equation violating the uncertainty principle by treating carriers as point charges with simultaneously well defined position and velocity.

3.2.3 Semi-Classical Equations of Motion

It is evident that, through the drift terms, the distribution function depends on the carrier dynamics. However, the dynamics of classical electrons in a semiconductor material is a many body problem and, for all intents and purposes, impossible to solve directly. In the independent electron approximation, electron-electron interactions are ignored and instead only the interaction between an electron and the crystal lattice is considered. Quantum mechanical solutions of the electron wavefunction that considers only the periodic potential of an infinite perfect lattice exist in the form of Bloch waves. These constitute a plane wave multiplied by a function with the periodicity of the lattice [98].

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) \quad (3.3)$$

where

$$u_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) \quad (3.4)$$

\mathbf{R} being a lattice vector. These functions quantize the allowed values of the electron wave-vector, \mathbf{k} , and associate with each an energy, $\varepsilon(\mathbf{k})$, which together define the dispersion relation, or band structure. The subscript n identifies different solutions of $\varepsilon(\mathbf{k})$, defining alternate bands, and is termed the band index. For the rest of this consideration the band index shall be dropped, assuming only one band. The dynamics of a single electron can now be described semi-classically by the equations of motion [99]

$$\frac{d\mathbf{r}}{dt} = \frac{1}{\hbar} \nabla_{\mathbf{k}} H \quad (3.5)$$

$$\frac{d\mathbf{k}}{dt} = -\frac{1}{\hbar} \nabla_{\mathbf{r}} H \quad (3.6)$$

where

$$H(\mathbf{r}, \mathbf{k}) = \varepsilon(\mathbf{k}) + eV(\mathbf{r}) \quad (3.7)$$

is the Hamiltonian, including the electrostatic potential $V(\mathbf{r})$.

At this point it becomes convenient to introduce the electron effective mass which is a useful concept in the analytical description of electron dynamics and makes analogous the familiar free electron dynamics. Associated with the electron wavevector \mathbf{k} is the electron's crystal momentum \mathbf{P} . These are related through the de Broglie relationship which can in turn be presented as the product of a mass-like term m^* and the velocity \mathbf{v}

$$\mathbf{P} = \hbar\mathbf{k} = m^* \mathbf{v} \quad (3.8)$$

m^* differs from the free electron mass m_e and is termed the effective mass. In general m^* is not constant but a function of the wavevector, thus its value varies along the band

structure. An applied force \mathbf{F} results in a change in the electron's crystal momentum $\hbar\mathbf{k}$ according to Newton's laws. Considering the components of the acceleration, Newton's law may be written as [75]

$$a_i = \frac{dv_i}{dt} = \frac{d}{dt} \frac{1}{\hbar} \frac{\partial \epsilon}{\partial k_i} = \sum_j \frac{1}{\hbar^2} \frac{\partial^2 \epsilon}{\partial k_i \partial k_j} \frac{d(\hbar k_j)}{dt} \quad (3.9)$$

or

$$a_i = \sum_j \left(\frac{1}{m^*} \right)_{ij} F_j \quad (3.10)$$

Where $\left(\frac{1}{m^*} \right)_{ij}$ is the inverse effective mass, written in general as a second rank tensor

$$\left(\frac{1}{m^*} \right)_{ij} = \frac{1}{\hbar^2} \frac{\partial^2 \epsilon}{\partial k_i \partial k_j} \quad (3.11)$$

This can be seen to be the velocity, as given by equation 3.5, divided by the momentum, given by equation 3.8, analogous to the free electron case. The effective mass is thus defined entirely by the band structure, which is discussed in section 3.3.1, and contains no new information.

3.2.4 Semi-Classical BTE

The distribution function can now be written in terms of \mathbf{r} , \mathbf{k} and t by substituting equations 3.5 and 3.6 into 3.2, resulting in the semi-classical BTE

$$\frac{\partial g(\mathbf{r}, \mathbf{k}, t)}{\partial t} + \frac{1}{\hbar} \nabla_{\mathbf{k}} \epsilon(\mathbf{k}) \cdot \nabla_{\mathbf{r}} g - \frac{1}{\hbar} \nabla_{\mathbf{r}} eV(\mathbf{r}, t) \cdot \nabla_{\mathbf{k}} g = \left(\frac{\partial g}{\partial t} \right)_{Coll} \quad (3.12)$$

The above equation includes a general spatially and temporally varying potential $V(\mathbf{r}, t)$ as would be required in a device simulation and/or transient problem, along with a general dispersion relation, $\epsilon(\mathbf{k})$, dependent upon the semiconductor material. This interpretation

of electron dynamics is again classical, assuming simultaneously knowing the position and velocity. The collisional term on the right hand side expresses the rate of change of the distribution function due to all available scattering processes. The scattering can either reduce the value of the distribution function at $(\mathbf{r}, \mathbf{k}, t)$ by instantaneously changing the \mathbf{k} -vector from \mathbf{k} to \mathbf{k}' (Out-Scattering) or oppositely can increase the value by changing \mathbf{k}' to \mathbf{k} (In-Scattering). In this approach, the quantum mechanical nature of the real problem is only taken into account in two ways. First through the use of the band structure to describe the interaction of an electron with the crystal lattice, and second in the evaluation of scattering rates through perturbation theory using Fermi's Golden Rule. These scattering rates, discussed in section 3.3.2, are also dependent upon the band structure. Thus, this a semi-classical treatment.

Equation 3.12 assumes no sources or sinks of particles which would otherwise effect the distribution function. If included, this would be accounted for by a term similar to the collisional term. Such terms represent the cumulative influence of all possible scattering, generation and recombination mechanisms. In Monte Carlo simulations, generation and recombination are typically not included as they occur on comparatively much larger timescales than that of the scattering mechanisms and play very little role in MOSFETs which are majority carrier devices [100]. The solution of equation 3.12 via the Monte Carlo method is described in the next section.

3.3 Monte Carlo Solution of the BTE

As previously stated, Monte Carlo is a general mathematical tool and even applied to charge transport in semiconductors the method can take different forms. As applied here, the MC method is conceptually a very simple iterative process that directly simulates the physical motion of carriers. The distribution function is evaluated by the repeated propa-

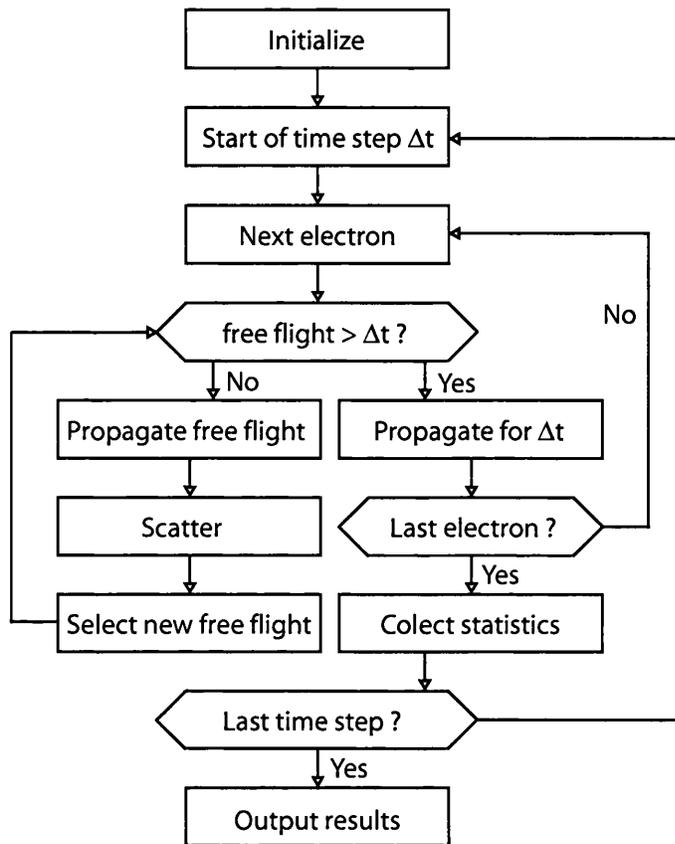


Figure 3.1: Basic flow diagram of Monte Carlo program suitable for simulation of bulk material

gation and self-consistent scattering of carriers, representing the drift and collision terms in the BTE, until the solution converges. This is in contrast to purely numerical algorithms with a less direct analogue to the physical system that are designed to discretise and solve the transport equation directly, though both methods are fundamentally linked [96], or statistically weighted methods that distort the real physical process to better resolve rare events [101].

In the practical implementation of the MC approach, the carriers are propagated in the local field according to the semi-classical equations of motion, 3.5 and 3.6, over a randomly generated free-flight time. At the end of the free-flight, carriers are scattered by

a randomly selected scattering mechanism into a new state and a new period of free flight follows. By gathering statistics on the state of each electron as the simulation progresses, we obtain an estimate of the distribution function. The final solution is obtained when averaging over a sufficiently large number of iterations gives convergence to within a desired accuracy. The basic flow diagram of the MC simulator is shown in figure 3.1. Self consistency, if required, is obtained through additionally solving Poisson's equation before each propagation time step, the new field is then determined and used for the subsequent propagation step.

For steady-state, homogeneous problems, it is sufficient to simulate the propagation of only one electron for a long time and apply the ergodic principle to treat the time average as an ensemble average. In cases where transient behaviour is under investigation or where the problem is non-homogeneous, an ensemble of carriers must be used to obtain the average directly in the Ensemble Monte Carlo (EMC) method. For the purpose of this work the latter choice is necessary since, even in simulations of a simple bulk semiconductor sample, the investigation of random discrete dopants results in an inhomogeneous simulation domain. Self consistency may then be incorporated into EMC simulations to recover the dynamic transient response of a system, or the steady-state solution alone may be efficiently achieved using a 'frozen field' solution in an appropriate steady-state potential. The simulation of bulk semiconductor properties in absence of discrete doping represents a simple homogeneous problem where self consistency is not necessary. After discussing the choice of fundamental models that are employed within the simulator throughout this work, such bulk simulation serves to provide a suitably simple method of calibration and verification of the simulator.

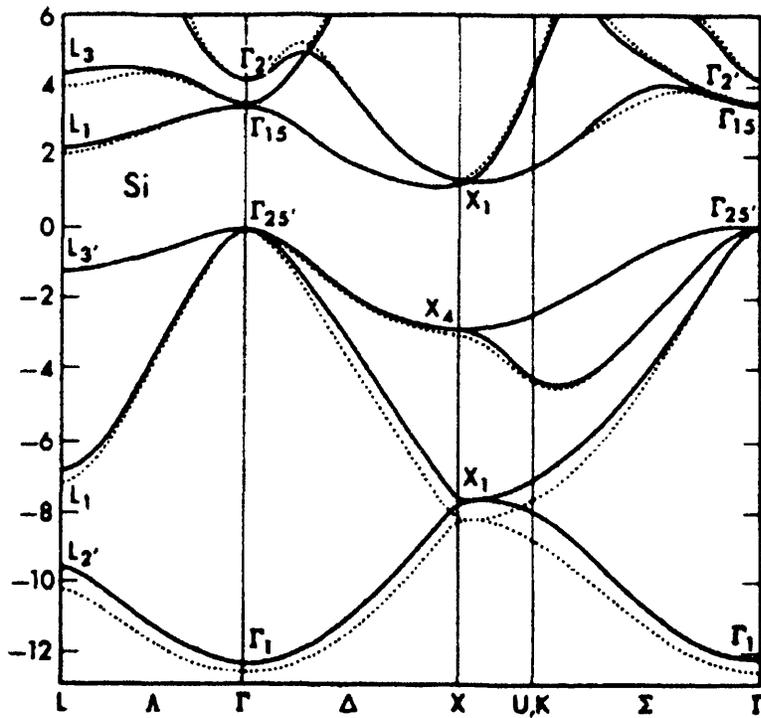


Figure 3.2: Silicon band structure calculated from non-local pseudopotential method (solid line) and local- pseudopotential method (dashed line) [103]

3.3.1 Band Structure

In order to propagate electrons and calculate scattering rates, a description of the band structure, $\epsilon(\mathbf{k})$, is required. Results from the theoretical calculations of the band structure of silicon are presented in figure 3.2. It can be seen that in general the band structure is very complicated. A simplification of the problem can be made by noting that in silicon the majority of electrons involved in transport at low fields reside near the conduction band minima, located at six equivalent valleys situated $0.85 \times \frac{2\pi}{a}$ from the the Γ point along the Δ symmetry line, near the X points in the Brillouin Zone (BZ), where a is the lattice constant [75, 102]. About this point the band may, to first approximation, be represented by a parabolic dispersion relation

$$\varepsilon(\mathbf{k}) = \gamma = \frac{\hbar^2 \mathbf{k}^2}{2m^*} \quad (3.13)$$

This is analogous to the free electron energy corresponding to a scalar effective mass, m^* . If during a simulation, electron energies exceed the range where this approximation holds, the range of applicability may be extended by the introduction of a fitting parameter, α , in the Kane-model [104]

$$\varepsilon(\mathbf{k})(1 + \alpha\varepsilon(\mathbf{k})) = \gamma = \frac{\hbar^2 \mathbf{k}^2}{2m^*} \quad (3.14)$$

where the energy is explicitly given by

$$\varepsilon(\mathbf{k}) = \frac{2\gamma}{1 + \sqrt{1 + 4\alpha\gamma}} \quad (3.15)$$

Whereas in the parabolic approximation the mass-like term, expressed solely by the scalar effective mass, is a constant, in the non-parabolic approximation it is an energy dependent quantity equal to the parabolic case only at the band minima. The fact that in general the effective mass is a second rank tensor (section 3.2.3) highlights a short-falling of both these analytical models. In equations 3.13 and 3.14, the energy is dependent on the magnitude of the wave vector alone and thus gives rise to spherical equi-energy surfaces in \mathbf{k} -space. In the actual band structure of silicon, the equi-energy surfaces around the band minima are ellipsoidal in \mathbf{k} -space [105], introducing anisotropy into the electron dynamics which is seen in experiment [106]. This can be taken into account using an analytical approximation with γ expressed as

$$\gamma = \frac{\hbar^2}{2} \left(\frac{k_x^2}{m_l^*} + \frac{k_y^2 + k_z^2}{m_t^*} \right) \quad (3.16)$$

where the effective mass is separated into longitudinal and transverse components with respect to the axes of the equi-energy ellipsoid, assumed here to be aligned along the

x-direction in k-space. Combinations of equations 3.13 - 3.16 give either spherical or ellipsoidal variants of the parabolic or non-parabolic band minima. Such approximations are increasingly less valid at high energies, and if accuracy in this regime is required then more detailed full band calculations may be employed [107–109].

Band structures are important not only for determining the dynamics of carriers but also for accurately calculating the scattering rates mainly through the influence of the density of states, an important, band dependent, property. Although full band solutions give a more realistic account of the band structure at higher energies, analytic approximations are easier and faster to employ and describe the dynamics of the majority of carriers well. Even within full band simulations, analytic approximations may be employed at the band minimum to avoid interpolation errors caused by resolving the band structure too coarsely [108, 109]. For these reasons, analytic bands are still widely employed in practical device simulations while full-band simulations are used when studying effects sensitive to the high energy electron distribution. For the purposes of this work, the accurate details of the high energy electron distribution are unimportant compared with the transport behaviour of the majority of electrons. For this reason the analytical spherical, non-parabolic band approximation is employed. This reproduces the conduction band accurately to about 200meV [110] and roughly approximates the density of states up to 2eV [111]. The disparity between the longitudinal and transverse effective masses is taken into account by using an averaged effective mass combining the longitudinal and transverse masses [75]

$$m^* = (m_l m_t^2)^{\frac{1}{3}} \quad (3.17)$$

This does not reproduce the experimentally observed anisotropic transport but gives a sufficiently accurate average description. With this approximation the semi-classical equa-

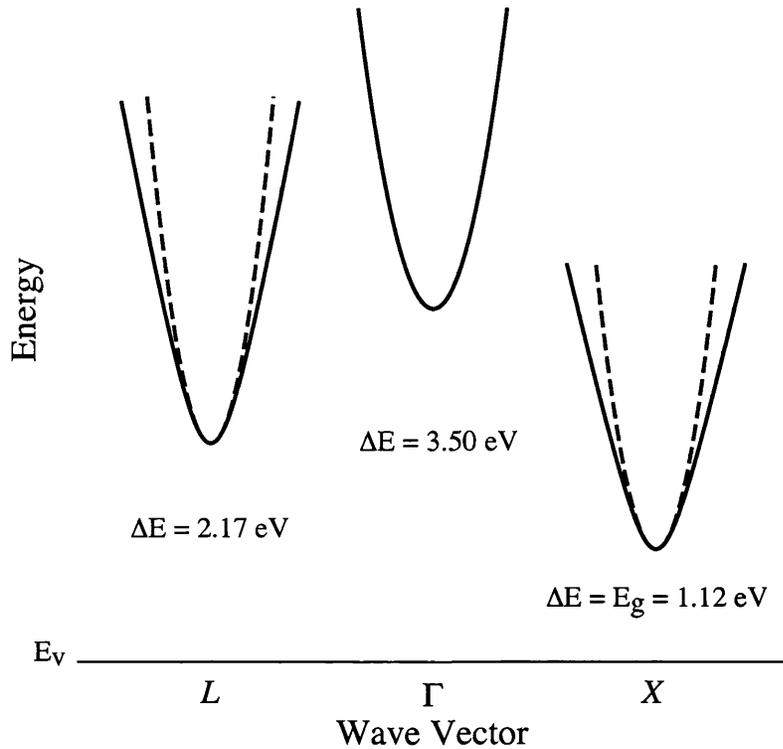


Figure 3.3: Diagram of the analytic band structure showing the three conduction band minima considered. Parabolic energy bands are shown as dashed lines for the X and L minima for comparison to the non-parabolic bands used. The separation in energy from the top of the valence band is also shown.

tion of motion are written as

$$\frac{d\mathbf{k}}{dt} = -\frac{e}{\hbar} \nabla_r V \quad (3.18)$$

$$\frac{d\mathbf{r}}{dt} = \frac{\hbar \mathbf{k}}{m^* (1 + 2\alpha \epsilon(\mathbf{k}))} \quad (3.19)$$

and the energy-dependent mass-like term in the velocity, $m^* (1 + 2\alpha \epsilon(\mathbf{k}))$, is present in equation 3.19 for $\alpha \neq 0$.

Multiple conduction band valleys are represented by individual non-parabolic 'bands' separated in energy as depicted in figure 3.3. The effect of the non-parabolicity factor

α in lowering the energy is shown for the X and L valleys in the same figure. The Γ valley is assumed parabolic and spherically symmetric. As previously stated, in silicon at moderate fields nearly all carriers reside in the X valley. The higher L and Γ valleys could be included for completeness to allow inter-valley transitions, but they play little role in this work.

3.3.2 Scattering Mechanisms and Rates

Having established the treatment of the electron dynamics governing the drift terms of the BTE, we now turn our attention to the collision term that balances the equation. As previously stated, this term comprises all mechanisms, other than the electric field, responsible for transferring a carrier from state \mathbf{k} to a new state \mathbf{k}' . Such collisions are elastic if $\varepsilon(\mathbf{k}) = \varepsilon(\mathbf{k}')$ and can only serve to randomize the ensemble momentum. Ionized impurity scattering is an example of an elastic collision and such mechanisms alone can not relax the energy of carriers. In order that carriers could be able to equilibrate with the crystal lattice, they must gain or lose energy. By the emission or absorption of phonons of energy $\varepsilon_p = \hbar\omega_p$, carriers change their momentum and decrease or increase their energy. The rate at which carriers suffer a collision depends upon the specific mechanisms at work and the carrier energy. This rate must be established in order to both determine the period of ballistic motion between collisions and the scattering mechanism responsible for terminating it.

Phonon scattering is one of the most important mechanisms in bulk materials and dominates transport properties at higher energies. Impurity scattering is another important mechanism and dominates at high doping concentrations and at low energy. Additional scattering mechanisms that may be considered are electron-electron and electron-plasmon interactions. These interactions, between either a pair of electrons or between

an electron and the surrounding ensemble, redistribute momentum and energy and are important mechanisms in shaping the high energy electron distribution. In devices, the presence of heterojunctions or other boundaries imposes further scattering. In particular surface roughness scattering from the gate insulator interface in a FET channel limits mobility by re-distributing carrier momentum upon impact. For the purposes of the work carried out here, only the dominant phonon scattering mechanisms are considered. This is justified by the aim to reproduce the dynamics of the majority of electrons without necessarily simulating the high energy tail. A method for treating ionized impurity scattering through the real space trajectories of the carriers will be introduced later, avoiding the use of scattering rates. In all cases, the scattering mechanisms are considered weak enough to be treated by means of first order perturbation theory as described in [75]. The scattering mechanisms included in this work are described next.

3.3.2.1 Acoustic phonons

Phonons arise from the collective oscillations of the atoms in the crystal lattice that distort the periodic crystal potential and alter the electron dynamics. They are directly associated with the quantized vibrational modes, and the corresponding wavevector, set up within the crystal lattice, and are characterised by their energy as a function of that wavevector. An illustration of the simplified phonon dispersion relation considered within this work is shown in figure 3.4. The acoustic phonon branch specifically refers to the collective oscillations of neighbouring atoms in the same direction. The effect that the displacement of each atom has on the deviation of the periodic potential from its equilibrium value is characterised by the acoustic deformation potential, Ξ_{ac} . The value of this constant governs the strength of the acoustic phonon interaction with the electrons and is usually used as a calibration parameter. Rigorous calculations of the scattering rate also require the evaluation of the overlap integral which weights the probability of scattering between

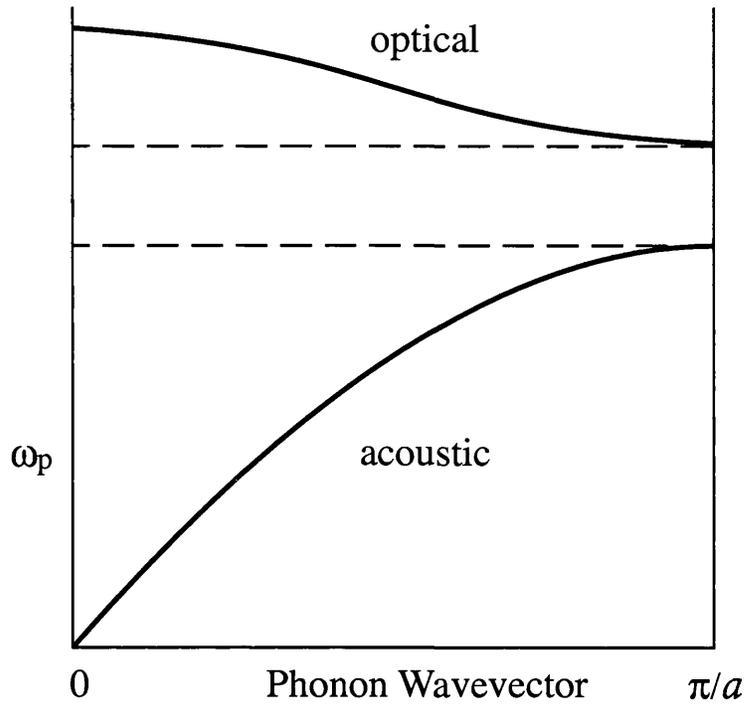


Figure 3.4: Illustration of the simple phonon dispersion approximation used within this work (reproduced from [112]). Only one acoustic and one optical branch is considered, neglecting transverse and longitudinal modes.

specific states. Here, following the discussion in [75], the overlap integral is assumed a constant and incorporated in the deformation potential. Scattering from acoustic phonons is treated as an inelastic, anisotropic process with the phonon dispersion curve as given in [108].

3.3.2.2 Optical phonons

Optical phonons differ from acoustic phonons in that neighbouring atoms oscillate in opposite directions. The scattering from an optical phonon mechanism is treated in a manner similar to the acoustic case but, for simplicity, the phonon dispersion curve for each mechanism is approximated as a constant. This approximation is based on the assumption that

Phonon Mechanism	θ_{g1}	θ_{g2}	θ_{g3}	θ_{f1}	θ_{f2}	θ_{f3}
Phonon Energy (meV)	12.06	18.53	62.04	18.96	47.40	59.03
Phonon Temperature (K)	140	215	720	220	550	685

Table 3.2: Phonon energies and associated phonon temperature for the considered optical phonons in silicon

each optical mechanism only involves a small range in phonon wave vectors. Justification of a constant energy for each phonon mechanism is then seen in the small variation of phonon frequency with wavelength in the optical branch (see figure 3.4). Scattering from optical phonons is isotropic and an important process for exchanging energy. We consider both intra-valley and inter-valley transitions due to optical phonons, of which the inter-valley scattering is further classified into transitions between equivalent valleys lying parallel (*g-type*) and perpendicular (*f-type*) to each other. Three such phonons for each process, *g-type* and *f-type*, are accounted for, with energies characterised by the phonon's energy, ϵ_{op} , or equivalent temperature, θ_{op} . The equivalent temperature is related to the phonon energy by $\epsilon_{op} = \hbar\omega_{op} = k_b\theta_{op}$ and is listed for each phonon in table 3.2 [109]. The strength of each interaction is governed by a coupling constant in which the overlap integral is again assumed constant and incorporated. As in the case of acoustic phonons, the coupling constants need to be calibrated.

3.3.2.3 Ionised Impurity Scattering

Although a treatment of ionized impurity scattering directly from the unique arrangement of impurity atoms will be developed, removing the need to describe the interaction via a scattering rate, in some simulations we will make use of the scattering rate description. Detailed discussion of the impurity scattering rate models and why the use of impurity scattering rates is unsuitable for simulation of nano-scale devices are deferred until Chap-

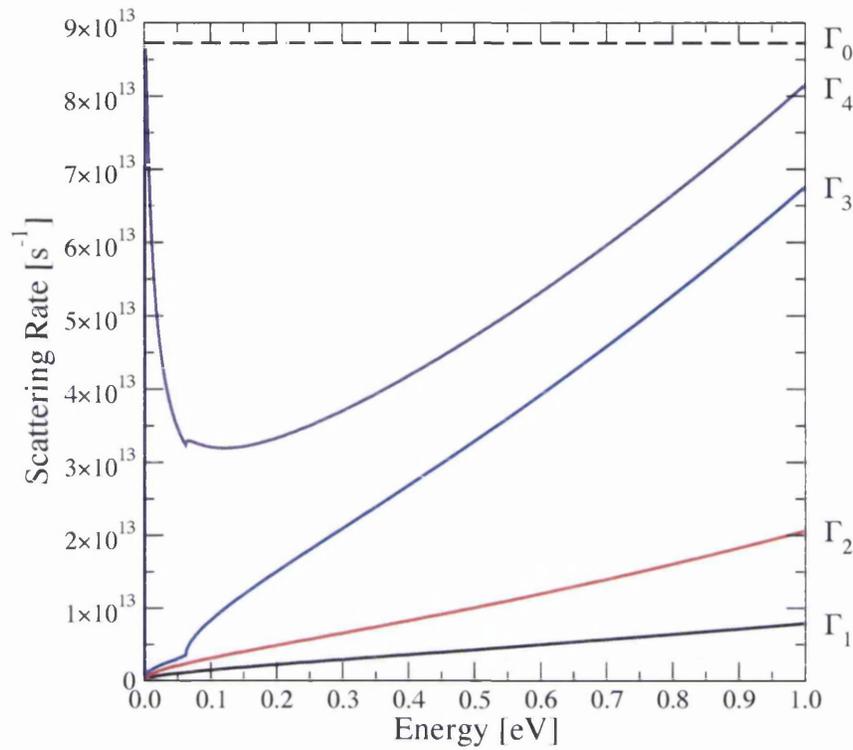


Figure 3.5: Calculated scattering rate in silicon at 300K

ter 4.

3.3.2.4 Total scattering rate

The total scattering rate is obtained summing the individual scattering rates, assuming the mechanisms are independent. A series of cumulative scattering rates, Γ_i , for the X -valley in silicon at 300K are shown in figure 3.5. Γ_1 is the energy dependent scattering rate for the acoustic absorption process alone, while Γ_2 additionally includes the acoustic emission rate. Γ_3 further includes all g -type mechanisms and shows the onset of the g_3 phonon emission mechanism at around 0.05eV . The total energy dependent scattering rate including all phonon mechanisms and ionized impurity scattering is shown as Γ_4 and finally Γ_0 marks the energy independent maximum scattering rate over the energy range

shown. The total scattering rate $\Gamma(\varepsilon)$, represented as Γ_4 in figure 3.5, defines the mean free flight time according to

$$\tau(\varepsilon) = \frac{1}{\Gamma(\varepsilon)} \quad (3.20)$$

and is used to determine the individual electron free flight times.

3.3.3 Free Flight

In the MC simulation cycle, each electron is subject to a randomly generated period of free flight where it is allowed to propagate in the local electric field following the semi-classical equations of motion, equations 3.5 and 3.6. The free flight is terminated by a scattering event, which is randomly selected from the available mechanisms consistent with their individual probabilities at a given electron energy, before a new free flight time is generated. Generation of free flight times must therefore satisfy the correct probability distribution as related to the total scattering rate.

The definition of the scattering rate is such that the probability P of an electron being scattered during time Δt is given by [105]

$$P\Delta t = \Gamma(\mathbf{k}(t))\Delta t \quad (3.21)$$

where $\Gamma(\mathbf{k}(t))$ (equivalent to $\Gamma(\varepsilon)$) is the energy dependent scattering rate at time t . Hence the probability that an electron has not undergone a collision since time $t = 0$ is given by

$$\exp\left[-\int_0^t \Gamma(\mathbf{k}(t')) dt'\right] \quad (3.22)$$

and consequently, the probability that an electron is subject to a scattering event a time t

after its last collision is

$$P(t) dt = \Gamma(\mathbf{k}(t)) \exp\left(-\int_0^t \Gamma(\mathbf{k}(t')) dt'\right) dt \quad (3.23)$$

The free flight times should be selected from this distribution. Unfortunately, a direct mapping of a uniform random deviate onto this distribution is complicated by the initial value and time dependence of $\Gamma(\mathbf{k}(t))$, specific to each electron's energy propagation (see figure 3.5). This problem is greatly simplified by the inclusion of a fictitious scattering mechanism, termed self-scattering, that does not alter the state of the electron but the inclusion of which increases the total scattering rate to a value independent of \mathbf{k} [113]. The total scattering rate, including self-scattering, becomes energy independent, allowing equation 3.23 to be re-written as

$$P(t) dt = \Gamma_0 \exp(-\Gamma_0 t) dt \quad (3.24)$$

The energy independent scattering rate Γ_0 is taken to be the maximum scattering rate evaluated over the energy range and scattering mechanisms considered, as is illustrated in figure 3.5. Using this formulation, a uniform random deviate, r , can be generated in the range $0 < r \leq 1$ and used to generate free flight times via the following expression

$$t_{ff} = -\ln(r)/\Gamma_0 \quad (3.25)$$

This is the most common way of selecting free-flight times and the method adopted hereafter. It has the benefit of being very easy to implement but results in the treatment of many fictitious self-scattering events at low energies. Efficiently identifying the scattering mechanism responsible for terminating a free flight helps minimise this overhead. Another method using a small, fixed, flight time and successively determining the probability

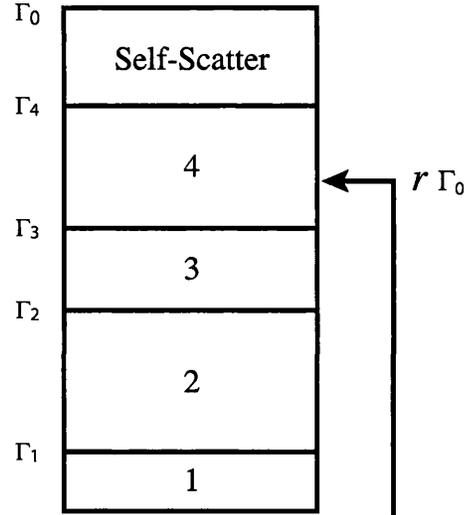


Figure 3.6: Selection of a scattering mechanism is made through comparison of a random fraction of the total rate with the cumulative sum of the individual rates. Here four scattering mechanisms along with the fictitious self-scattering are depicted.

of scattering after this time has been used and may provide an efficient alternative [108].

3.3.4 Selection of Scattering Event

Once a carrier has reached the end of its free-flight a scattering mechanism must be chosen from the mechanisms available, including self-scattering, according to their relative probabilities as a function of carrier energy. This is determined by comparing the product of a random number r , in the range $0 < r \leq 1$, and the energy independent scattering rate Γ_0 with the successive sums of the individual mechanism rates, Γ_i . The i^{th} mechanism is chosen when the condition

$$\sum_{n=1}^{i-1} \Gamma_n(\epsilon) \leq r\Gamma_0 < \sum_{n=1}^i \Gamma_n(\epsilon) \quad (3.26)$$

is met.

In order to avoid the computationally expensive repeated calculation of scattering rates, an efficient procedure following from the discussion of *fast self-scattering* in [114] is adopted. Here efficiency is achieved by tabulating the cumulative scattering rates of successive mechanisms over the energy range $[0, E_{max}]$ for each band. E_{max} should be large enough to encompass the range of expected energies during the simulation if excess computation of scattering rates is to be avoided. The energy range is divided into intervals, ΔE , at which the scattering rates are calculated. This division should be sufficiently fine in order to resolve the energy dependence of the scattering rates. A value of 1 meV is used throughout this work. With this table constructed, a mechanism is selected by again generating a random number, r , in the range $0 < r \leq 1$ and multiplying it by the energy independent scattering rate, Γ_0 . Selection is then made in the same way as 3.26, except comparison is made with stored values in a given energy interval. This is illustrated in figure 3.6.

3.3.5 Statistics Gathering

At intervals during the simulation, the state of the carrier/carriers must be sampled in order to gather statistics for evaluating the distribution function. In the simple one electron MC simulations, an ensemble average is inferred from the time average of the carrier state. This time average is evaluated by sampling the carriers position and momentum at the end of each free flight, prior to the scattering event [115]. These values are weighted by the free flight time, summed and divided by the total simulated time. Following this, the time average of a quantity A is defined as

$$\langle A \rangle_t = \frac{1}{T} \sum_{f=1}^{N_f} A_f t_f \quad (3.27)$$

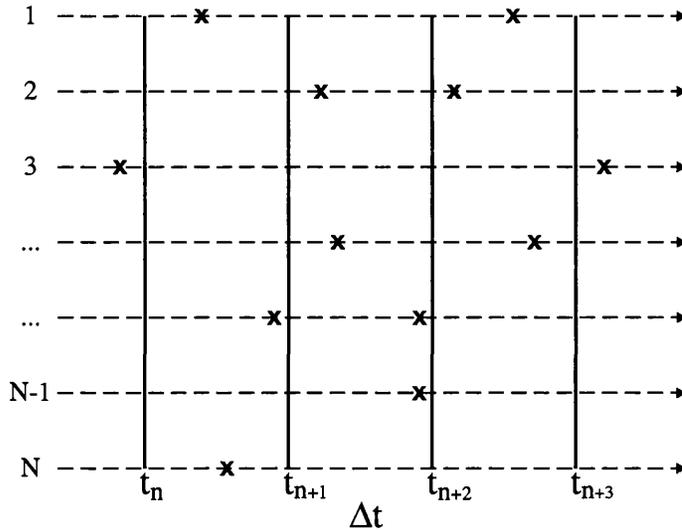


Figure 3.7: The state of carriers in an ensemble are sampled at regular intervals in time to produce synchronised averages. This is independent of the scattering events, marked by crosses, unlike the one electron case.

where N_f is the number of free flights during the time T and A_f is the value of the quantity of interest at the end of the free flight of duration t_f .

In EMC simulations the ensemble average of a quantity may be obtained directly from the simulated ensemble of carriers. This allows time averaged ensemble averages, or ensemble averaged time averages to be combined in various ways. To achieve this, the simulated time is divided into a series of short sub-histories of length Δt . All particles are propagated in time to the end of each sub history where the ensemble averages are calculated. This is termed the synchronous-ensemble and is illustrated in figure 3.7. Over the time interval Δt , time averages may be calculated by equation 3.27 and the ensemble average may then be obtained thereafter. This makes the best use of available information. Alternatively, the ensemble average may be approximated at the end of Δt by taking an average of instantaneous values. This is efficient to implement when performing self-consistent simulations (see section 3.5) where carriers are necessarily brought to the same

point in time to re-evaluate the electrostatic potential. For both cases a final average is taken over all averages separated by Δt to yield the time averaged approximation to the distribution function. In doing so an initial transient period is neglected and only the steady state simulation period is used to obtain the distribution function.

The last method is adopted here. While this may not make full use of the available information, the time step Δt used is much shorter than the mean free flight time and as such most time steps reflect a short period of ballistic motion. This approximates well the evaluation based on equation 3.27 which was predominantly used with larger time steps [116].

3.4 Bulk Transport in silicon

The discussions of band structure, semi-classical equations of motion, scattering rates and extracting results encompasses everything needed to analyse bulk properties of semiconductors in a homogeneous field. However, reliable simulations first require calibration of the phonon mechanisms to ensure the reproduction of experimentally observable results before meaningful predictions may be obtained. In bulk simulation, the steady-state distribution function can be obtained from either the one particle MC or EMC methods. Since the EMC approach must be adopted in the simulation of discrete dopants, it has been adopted from the beginning and is also used here during the calibration. Agreement of simulated bulk transport results with experiment provide validation of the approximations and parameters used. Bulk simulation is also used here to highlight the typical results obtainable with the EMC method.

T (K)	77	110	160	245	300
E ($kV\ cm^{-1}$)	0.0292	0.0580	0.1197	0.2571	0.4704
μ_{ToF} ($cm^2V^{-1}s^{-1}$)	20084	10115	4920	2355	1282
μ_{Ohm} ($cm^2V^{-1}s^{-1}$)	20540	11121	5431	2205	1423

Table 3.4: Experimental drift mobilities in silicon measured by time of flight technique for varying temperatures and electric fields [109]

3.4.1 Calibration

The acoustic phonon deformation potential, along with coupling constants for all optical phonon modes considered, need to be calibrated to reproduce experimentally observed properties. This was achieved through comparison with data for low-field electron mobilities in bulk Silicon. Experimental and theoretically calculated values of low-field mobility as functions of both temperature and field, taken from [109], are tabulated in table 3.4.

Initially, simulations were performed using a lattice temperature of 77K such that the optical phonon interactions are frozen out due to their low occupation. Under these conditions the acoustic deformation potential was adjusted to provide agreement with the tabulated mobility. Then, by increasing the lattice temperature and applied electric field, further mechanisms were successively introduced which best permitted the independent adjustment of each coupling constant to maintain agreement with mobility. The optical phonon energies and associated temperature, shown in table 3.2, indicates the temperature at which each phonon mechanism starts becoming statistically important compared to lower temperatures. This was used in conjunction with the simulated lattice temperature to decide which mechanisms required adjusting. The calibrated acoustic deformation potential and optical phonon coupling constants, along with other parameters relevant to the simulation, are given at the end of this chapter in tables 3.6 and 3.8.

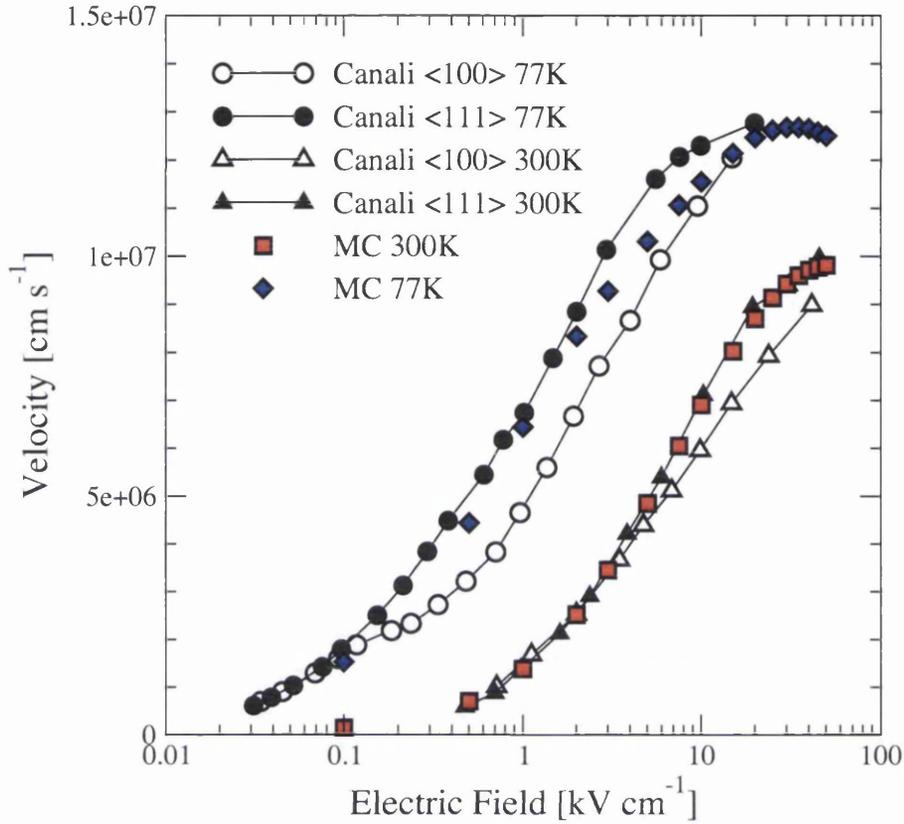


Figure 3.8: Ensemble Monte Carlo (MC) Simulated bulk velocity field curves in silicon at 77K and 300K compared with experimental values (Canali) [106]

3.4.2 Bulk Velocity-Field Characteristics

In an un-doped bulk silicon sample the scattering is dominated by the phonon processes and the ionized impurity scattering can be neglected. Therefore verification of the phonon calibration can be provided through the comparison of simulation results with experimental bulk velocity-field characteristics. In figure 3.8, simulation results are compared with experimental data for silicon at temperatures of 77K and 300K [106]. Experimental data clearly shows the anisotropic behaviour of transport associated with the ellipsoidal valley around the six equivalent band minima. As previously mentioned (section 3.3.1) this anisotropy is ignored in the spherical band approximation used here and is unobservable

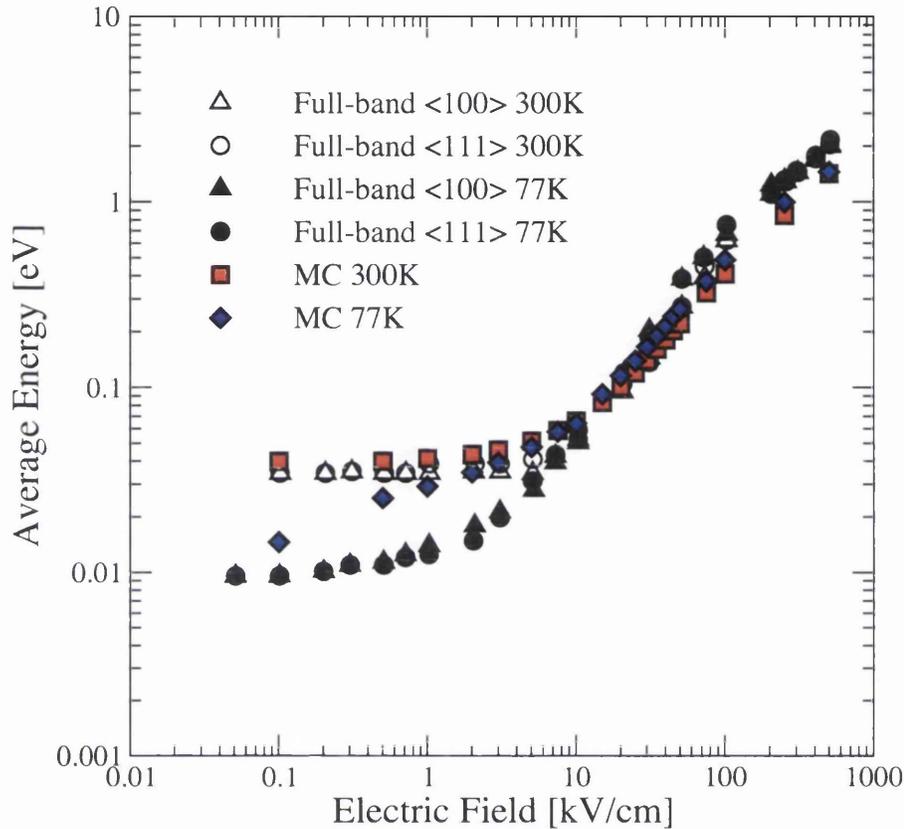


Figure 3.9: Simulated ensemble average energy as a function of applied field for this work (MC) in silicon at 77K and 300K compared to full band simulations [108]

in our simulation results. Inspection of the experimental velocity-field curve in figure 3.8 shows that the effect of anisotropy is small at low fields, and low energy, at 300K. All future simulations are carried out at 300K and at low field where possible. Since our intention is not to accurately model the high energy distribution, the analytic spherical band model is still a good approximation. As the anisotropy can not be reproduced with the adopted model, the phonon mechanisms are calibrated to reproduce an average result that lies between the results for the [100] and [111] directions.

3.4.3 Bulk Energy-Field Curve

Further validation of the simulation methodology is obtained by comparison of the average electron energy, as a function of electric field, with results obtained by full band simulation [108]. This is shown in figure 3.9, again for an un-doped sample, where the trend is reproduced well. The analytical band used here results in a slightly higher average energy at low fields and a lower energy at high fields compared with the full band simulation, a result consistent with that reported in [108].

3.4.4 Bulk Concentration Dependent Mobility

Including ionized impurity scattering in the simulation allows the doping concentration dependence of the mobility to be approximated. Figure 3.10 shows the simulated low-field concentration dependent mobility as a function of donor concentration compared with experimental values [117]. The simulated mobilities were obtained using the Brooks-Herring model of impurity scattering with an applied field of 1 kV cm^{-1} . The simulations reproduce the mobility well at low concentrations, where phonon interactions dominate, and the trend of mobility reduction at higher doping concentrations. The Brooks-Herring model, however, overestimates the mobility at high doping levels where the scattering process and electron dynamics become very complicated.

3.4.5 Transient Response

One of the most significant advantages of the Ensemble Monte Carlo approach over the faster Drift Diffusion (DD) approach in device simulation is the ability to simulate transient responses and non-local, non-equilibrium transport phenomena. This advantage is best seen in the MC ability to reproduce the electron velocity overshoot response to sudden application of a large electric field. This effect is increasingly more important in

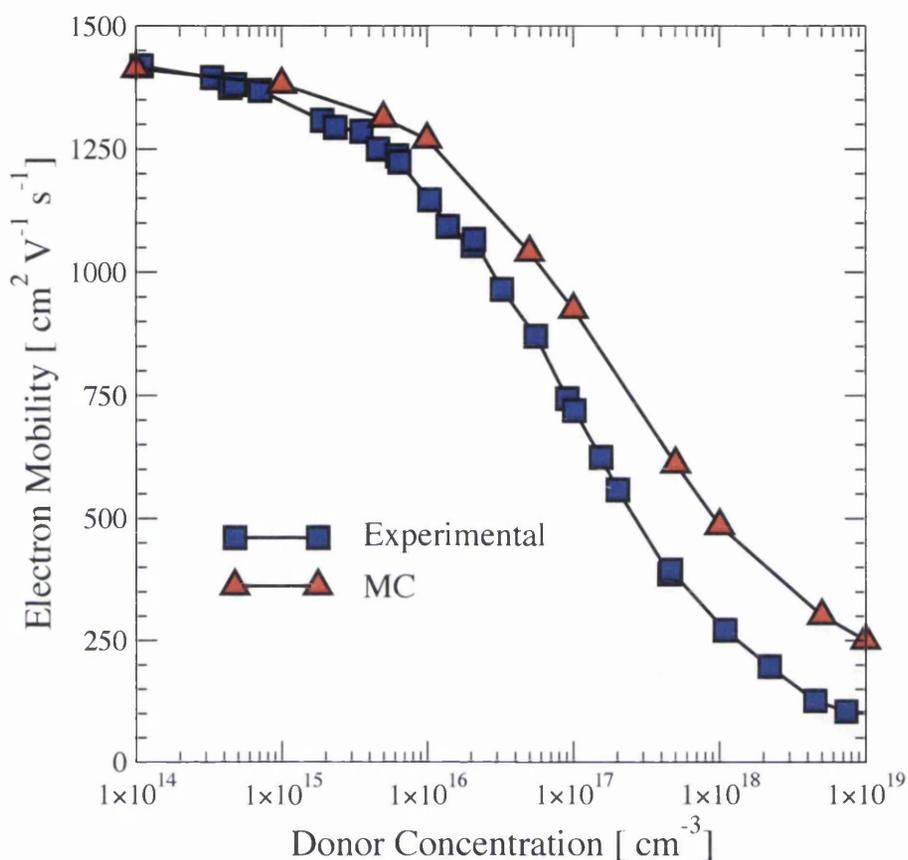


Figure 3.10: Simulated dependence of electron mobility upon doping concentration from bulk EMC simulation using Brooks-Herring scattering rate. Simulation results are compared to experimentally observed values [117]

short channel devices where substantial propagation through the channel may occur in overshoot regime. Such overshoot is unobservable in DD simulations where mobility is locally related to the electric field. More complicated energy transport models have some success in modelling the overshoot via phenomenological parameters and can extend the range of simulation validity, but the most accurate reproduction of the non-equilibrium transport phenomena is achieved by using the EMC solution. The ensemble averaged velocity response to the abrupt switching at $t = 0$ of a series of uniform applied electric fields is shown in figure 3.11. Electrons are initialized with random velocities resulting in

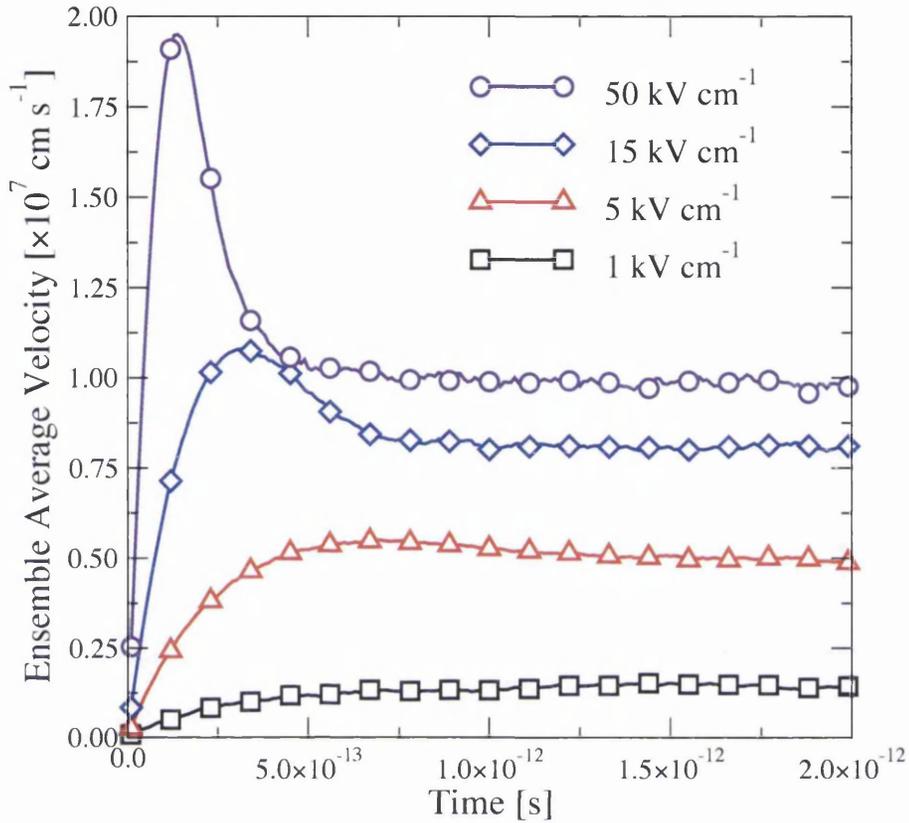


Figure 3.11: Simulated velocity overshoot response to a series of applied electric fields. At high applied fields the Monte Carlo simulation recovers a period of transport where electrons move faster than the steady state velocity, a process unobservable in Drift Diffusion simulation.

zero ensemble average velocity before the field is 'switched on'. At high field, electrons quickly gain energy, but loss of energy through phonon emission is delayed. This results in the initial peak in velocity which may be above the saturation velocity for silicon of around $1 \times 10^7 \text{ cm s}^{-1}$. At low field, electrons gain energy sufficiently slowly and the phonon mechanisms are able to control the velocity response up to the equilibrium value.

3.5 Device Simulation

Simulation of semiconductor devices follows from the bulk case with two important modifications. First, the electric field can no longer be considered constant. Instead the position dependent electric field resulting from the device structure, mobile charge density and applied biases must be calculated by solving the Poisson equation with appropriate boundary conditions and used during the electron dynamics phase. Second, the finite size of the device means that physical boundaries, representing reflective edges and absorbing metallic contacts, must also be accounted for in the carrier dynamics. Both modifications require the introduction of a simulation mesh upon which the device structure is mapped and Poisson's equation is discretized and solved. A number of possible solutions exist for the mesh construction, particle representation and force interpolation which can affect the accuracy and validity of simulations. The choices made here are outlined within the following section.

3.5.1 Solution Domain and Mesh Design

All simulated devices are mapped onto a rectangular simulation domain and discretized using a uniform 3D rectilinear mesh with N_x , N_y , N_z mesh points along the x , y , z axes respectively. The number of mesh points together with the device dimensions determines the cell widths, H_x , H_y , H_z . Material parameters are stored on the mesh point $P_{x,y,z}$ assuming a constant material between nodes over the range

$$[(P_x, P_y, P_z), (P_{x+1}, P_{y+1}, P_{z+1})] \quad (3.28)$$

This is in slight contrast to an alternative choice of mesh [99, 118] where the advocated mesh assigns material parameters to mesh nodes $P_{x,y,z}$ assuming uniform material over

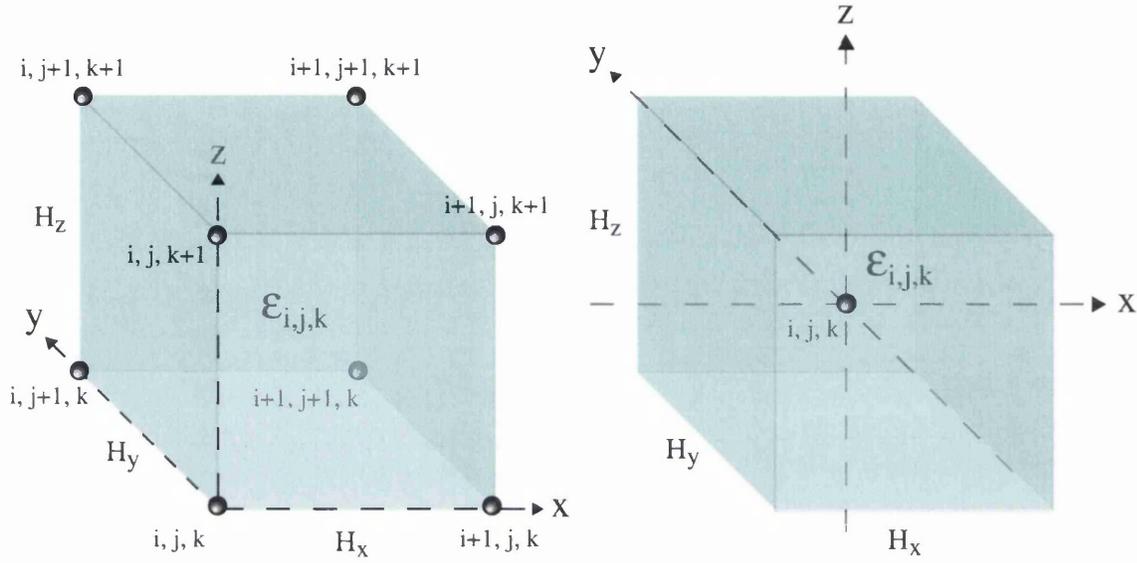


Figure 3.12: Two possible descriptions of the material associated with the device mesh. Material is either uniform between adjacent nodes (left) or uniform surrounding a node (right). The first description is adopted here.

the range

$$\left[\left(P_{x-\frac{1}{2}}, P_{y-\frac{1}{2}}, P_{z-\frac{1}{2}} \right), \left(P_{x+\frac{1}{2}}, P_{y+\frac{1}{2}}, P_{z+\frac{1}{2}} \right) \right] \quad (3.29)$$

These different schemes are illustrated for clarity in figure 3.12. The description chosen here makes the treatment of physical boundaries somewhat easier as they are naturally aligned along mesh lines. This choice of mesh is also consistent with the discretization of Poisson's equation which assumes a constant dielectric material between each node.

The mesh spacing defines the smallest size over which mesh assigned quantities may be resolved. For stability and accuracy it is important that this spacing is suitably chosen to be able to resolve the electrostatic potential throughout the device, and hence electric field, in which particles are propagated. The largest spacing is normally limited by the pre-estimated smallest value of the Debye length within the simulation, the characteristic length over which variations in potential decay in a neutral plasma. Competing

with this requirement is the need for efficient solutions of Poisson's equation that limits the smallest size that is practical to use. Further to this, it has been argued that using a fine mesh begins to resolve individual charge interactions, thus automatically including a component of ionized impurity or carrier-carrier scattering, resulting in 'double counting' of Coulomb scattering when including the corresponding scattering mechanisms in a stochastic process [108]. In the case of the *ab initio* treatment of ionized impurity scattering that shall be discussed further, both these restrictions upon the mesh can be relaxed in favour of a partially mesh-less force evaluation method, accompanied by the removal of corresponding ionized impurity scattering mechanisms and rates.

3.5.2 Charge Assignment

The Poisson equation is solved on the discretization mesh using an efficient Multigrid solver for given values of the charge density at each mesh point [119]. In order to approximate the charge density from the continuous particle positions, some form of interpolation to the mesh points must be used. Among such interpolation schemes are the Nearest Grid Point (NGP), Cloud In Cell (CIC) and Triangular Shaped Cloud (TSC), which are discussed in detail in [118]. An alternative scheme, the Nearest Element Centre (NEC), is reported in [120]. The method used herein is that of CIC which assigns charge over the surrounding mesh points weighted by the particle separation from these mesh points. This scheme has advantages in reducing fluctuations in the calculated electrostatic potential, and in removing the translational invariance within a cell that is a major problem of the lower order NGP scheme. Further smoothing and benefit could be obtained using the TSC scheme, however it is computationally more expensive than CIC and results in individual particles being represented on the mesh with a larger size compared to CIC. The NEC method assigns charge equally to the surrounding nodes, smoothing the charge

distribution but leading to invariance within a cell.

As all simulations in this work utilize a uniform mesh, the CIC assignment provides the best compromise between speed and accuracy. In this case, the charge density at any mesh point, $\rho_{x,y,z}$, is obtained as the sum of contributions from all N_e electrons in the simulation at that time, given by

$$\rho_{x,y,z} = \frac{1}{H_x H_y H_z} \sum_{i=1}^{N_e} q_i W(x_i - x) W(y_i - y) W(z_i - z) \quad (3.30)$$

where the weighting factors, W , are given by

$$W(x_i - x) = \begin{cases} 1 - \frac{|x_i - x|}{H_x} & |x_i - x| \leq H_x \\ 0 & \text{otherwise} \end{cases} \quad (3.31)$$

with x_i representing the x -position of the i^{th} electron and x the x -position of the mesh point. Equations similar to 3.31 apply for the y and z directions. In this case, assignment of charge density applies to the eight nodes surrounding a single particle with a linear weighting depending upon distance.

3.5.3 Force Evaluation and Interpolation

With the electrostatic potential $V_{x,y,z}$ calculated at each mesh point, the components of the electric field, E , at each mesh point are approximated using the following relations

$$E_X(x, y, z) = \frac{1}{2H_x} (V_{x-1,y,z} - V_{x+1,y,z}) \quad (3.32)$$

$$E_Y(x, y, z) = \frac{1}{2H_y} (V_{x,y-1,z} - V_{x,y+1,z}) \quad (3.33)$$

$$E_Z(x, y, z) = \frac{1}{2H_z} (V_{x,y,z-1} - V_{x,y,z+1}) \quad (3.34)$$

The subscripts X, Y, Z denote the Cartesian components of the field at mesh point $P_{x,y,z}$. The field at any point in the mesh is found by tri-linearly interpolating the field values at the surrounding nodes. This finite difference approximation and interpolation results in a continuous field throughout the mesh and results in zero self-force [118]. Self force is a problem in mesh based simulations and arises due to the finite size of the particle as represented in the mesh based solution of Poisson's equation. Evaluation of the mesh force and interpolation to the particle's position can result in a force contribution from the particle itself, un-physically accelerating it. Using a uniform mesh along with a force interpolation that is consistent with the charge assignment removes the problematic self-force [118].

For this choice of field calculation, significant errors may be introduced at the mesh points defining heterojunctions where there is a large change in the dielectric constant of the material. This may be avoided by calculating the field parallel to the edges of each mesh cell, obtained from the potential difference across the individual cell of constant material. If calculated this way, the components of the force are no longer continuous along the corresponding mesh directions. Such a system, however, suffers from self-forces that result in particles oscillating within mesh cells [118]. An alternative scheme of field calculation that attempts to account for the problem of a discontinuous dielectric by averaging the displacement field across the boundary is described in [108] and may prove a better choice for self-consistent device simulation.

For devices considered herein, only the semiconductor/insulator interface in a MOS-FET poses this problem. In this case, the field at the nodes defining this interface is determined from the dielectric alone. As the dielectric is charge neutral in all simulations, the electric field in the dielectric is constant and provides an accurate estimate of the magnitude of the field at the interface.

3.5.4 Field Adjusting Time Step

The solution of the Poisson equation is time consuming and therefore, during a self consistent simulation, it is desirable to solve it only as often as necessary. However, the time between successive solutions, the so called field adjusting time step, needs to be short enough so that the field can be considered constant during the particle propagation. This time is generally limited by the largest expected plasma frequency in the simulation, given by equation 3.35

$$\omega_p = \sqrt{\frac{ne^2}{\epsilon m^*}} \quad (3.35)$$

where n is the maximum electron concentration expected and is normally taken as equal to the the maximum doping concentration found within the device. The field adjusting time step is normally chosen to adequately sample this frequency [121, 122]. In the *ab initio* treatment of ionized impurity scattering discussed later, it is found that the major consideration limiting the time step is the accuracy to which the equations of motion for electrons are integrated. This imposes more severe restrictions on the time step compared to restrictions associated with the plasma frequency. For all device simulations the field adjusting time step is chosen as $10^{-16}s$.

3.5.5 Boundary Conditions

The solution of Poisson's equation requires the application of adequate boundary conditions. Dirichlet boundary conditions, fixing the value of the potential at specific mesh points, are used to represent the applied potential at device contacts. The remaining boundaries apply the Neumann boundary condition, forcing the gradient of the potential to be zero perpendicular to the edge of the simulation domain. These are reflected in the electron dynamics as contact regions and specularly reflective boundaries respectively.

Treatment of contacts within MC simulation remains a troublesome affair due to the

lack of understanding of the physical processes involved at the contact/semiconductor interface. In the simulation process, the contact is treated as closely as possible as an ideal ohmic contact. This idealized scenario defines a small region adjacent to the Dirichlet boundary throughout which charge neutrality is maintained. The contacts are defined in highly doped regions to ensure that little potential is dropped across them allowing a diffusive supply of carriers in and out of the region. Maintaining the region as charge neutral requires the removal of electrons that cross the contact out of the simulation domain and the injection of electrons through this contact if too few are present. An excess of electrons is allowed to disperse naturally through the contact, either being removed or propagating into the active region of the device. Injection of electrons, if required, is performed after the removal of electrons at the end of each propagation time step. The initial \mathbf{k} -vector of each injected electron is formed from a component perpendicular to the contact plane which is selected from a velocity-weighted Maxwellian distribution, and two components parallel to the contact plane both selected from a Gaussian distribution. This has been shown to produce better agreement with the behaviour of an ideal Ohmic contact as compared with the use of three Gaussian components [123, 124]. The initial position of the injected carriers is chosen to best satisfy charge neutrality at every point in the contact. In 3D simulation this can be a very hard to realise owing to the low number of particles per mesh cell. This is reflected in a rapidly varying potential dominated by individual charges, and ultimately means that special care must be taken regarding the initial position of newly injected electrons.

The requirement for charge neutrality determines the overall number of electrons in the contact region. In device simulation this involves balancing the electron charge with the charge of acceptors, donors and holes and in doing this we assume no intrinsic electron contribution. This is a valid assumption since the intrinsic concentration of electrons is orders of magnitude less than that required to balance the net doping charge. To

accurately position the electrons, the net charge density associated with each mesh cell describing the contact is evaluated as an average of the density at the mesh points surrounding the cell. Summing over all such cells within the contact results in a total net charge for the contact which is then multiplied by a random number between 0 and 1. This results in a target value for the net charge, in a process analogous to the selection of a scattering mechanism using the total scattering rate. Repeating the summation process and stopping when the target value is met selects the appropriate cell. Thus a cell into which the next electron is injected can be randomly selected with a probability weighted by the charge distribution at that time. An electron is then placed randomly within the chosen cell and the charge density updated to account for its presence. This process is repeated until all electrons required have been positioned. This procedure proves to be a very simple and effective way of ensuring that electrons are accurately distributed in space, helping maintain a stable Ohmic contact.

3.5.6 Current Calculation

For the purposes of this work we employ a simple method for estimating terminal currents. This consists of counting the net number of particles crossing the contact and performing a 5 point based numerical differentiation to obtain an average of the time rate of change. This is a very simple method but suffers from statistical noise in simulations with few particles, demanding longer simulations to improve current estimations. The Ramo-Shockley method for determining currents is more sophisticated and results in a more accurate evaluation of current [125]. Should additional accuracy be demanded, the implementation of the Ramo-Shockley current evaluation may be included.

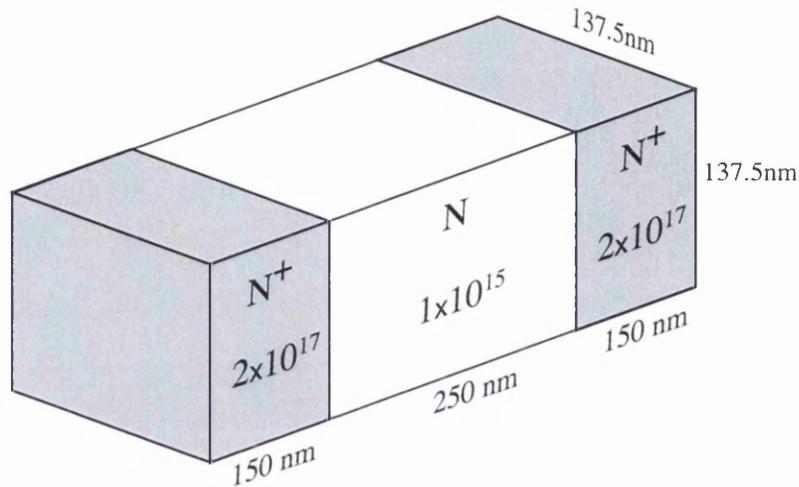


Figure 3.13: Diagrammatic representation of the N-I-N device, showing the dimensions in 3D and the doping concentrations throughout.

3.6 Example: Self Consistent N-I-N Diode

As an example, the simulation of a large $N-I-N$ diode using the developed 3D EMC simulator, where ionized impurity scattering is treated via the Brooks-Herring scattering rate model (see Chapter 4) is presented. The device structure is illustrated in figure 3.13 [99]. The device measures $550 \text{ nm} \times 137.5 \text{ nm} \times 137.5 \text{ nm}$ in the x , y and z directions respectively, and the discretization consists of $129 \times 33 \times 33$ mesh points. The external regions are doped at a concentration of $2 \times 10^{17} \text{ cm}^{-3}$, while the central, lower doped, region has a concentration of $1 \times 10^{15} \text{ cm}^{-3}$. Poisson's equation is solved every $1 \times 10^{-16} \text{ s}$ and the simulation consists of 70,000 such time steps (simulating 7 ps). All averages are calculated after an initial transient period of 20,000 time steps (2 ps) which is large enough not to affect the final results. The time step adopted is smaller than required to sample the plasma frequency at these doping levels [121], but is used to be consistent with later simulations where numerical accuracy is required. The simulation used an ensemble

of 40,000 particles.

3.6.1 Electrostatic potential and electron concentration

The simulated time averaged potential is shown as the full 3D plot in figure 4.28 and a 1D averaged profile from the same data in figure 3.15. The corresponding electron concentration is also shown using similar 3D and 1D plots in figures 3.16 and 3.17 respectively. The results reproduce the expected behaviour when compared with previous simulation studies [99, 126], clearly showing the depletion regions at the $N - I$ junctions, with most of the depletion region occurring at the reverse biased drain end. Good maintenance of charge neutrality in the contacts is also achieved.

3.6.2 Electron energy and velocity

The 3D and 1D plots of the average electron energy are shown in figures 3.18 and 3.19 respectively. Similarly plots of the velocity distributions are shown in figures 3.20 and 3.21. The energy steadily rises in the intrinsic region, starting from the thermal source energy and reaching a maximum of nearly 170 meV before rapidly thermalizing in the drain. On the other hand, the velocity quickly rises and reaches the saturation velocity, $1 \times 10^7 \text{ cm s}^{-1}$ before returning to a constant lower value in the drain.

The 3D results suffer from higher spatial noise compared to 1D or 2D simulations because of the low ratio of simulated particles to mesh cells. It becomes impractical to simulate enough particles to smooth the statistics. In this example there is 1 particle for every 3 mesh cells. This simple device geometry allows simulation in 1 dimension and therefore allows averaging over the y and z directions to produce the 1D profiles which have far less noise. In more complicated device geometries however, this will not be possible and is recognised as a limiting factor for efficient simulation.

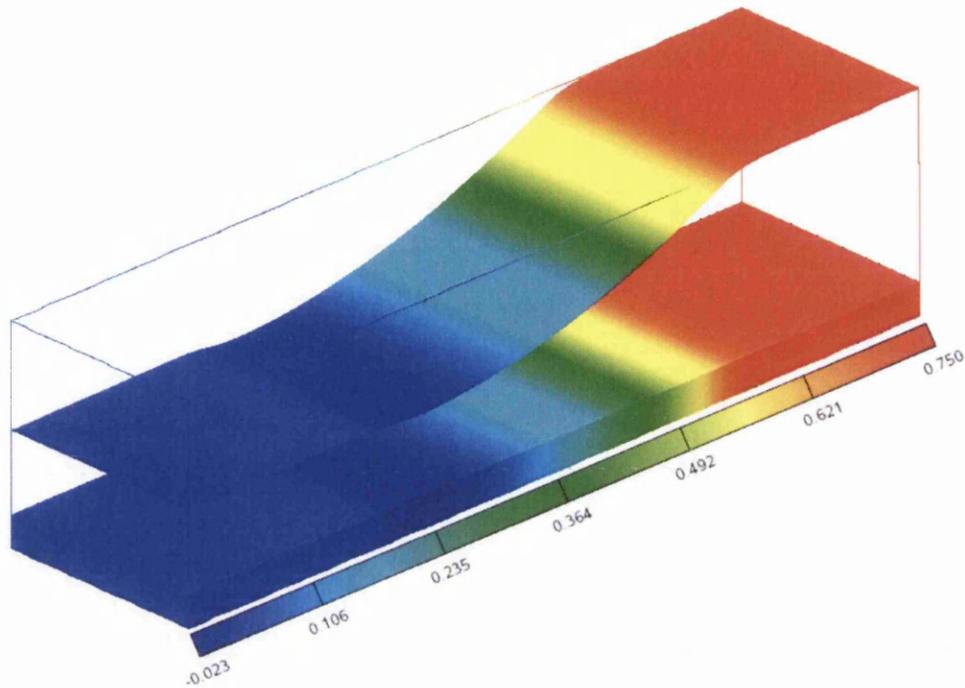


Figure 3.14: 3D distribution and profile of the time averaged electrostatic potential, in Volts, from EMC simulation

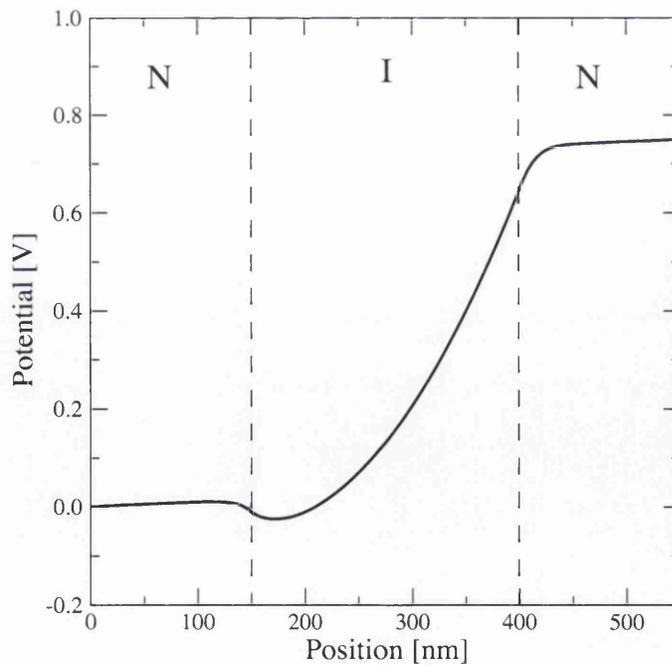


Figure 3.15: 1D profile of the electrostatic potential shown in figure 4.28 having averaged over the y and z directions. The boundaries between the N I N regions are marked by the dashed lines.

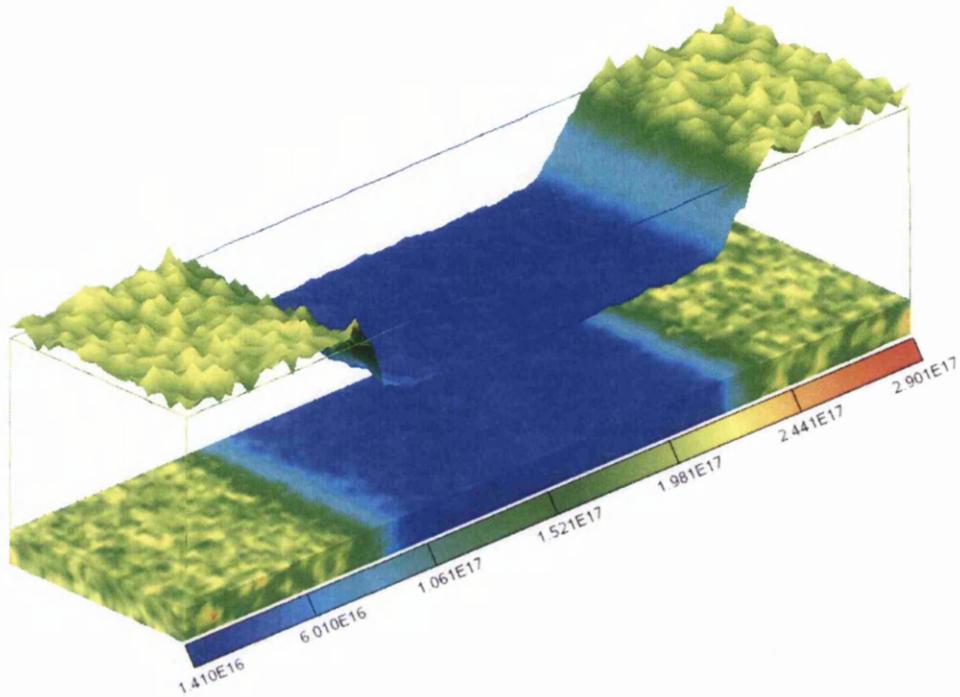


Figure 3.16: 3D distribution and profile of the time averaged electron concentration cm^{-3}

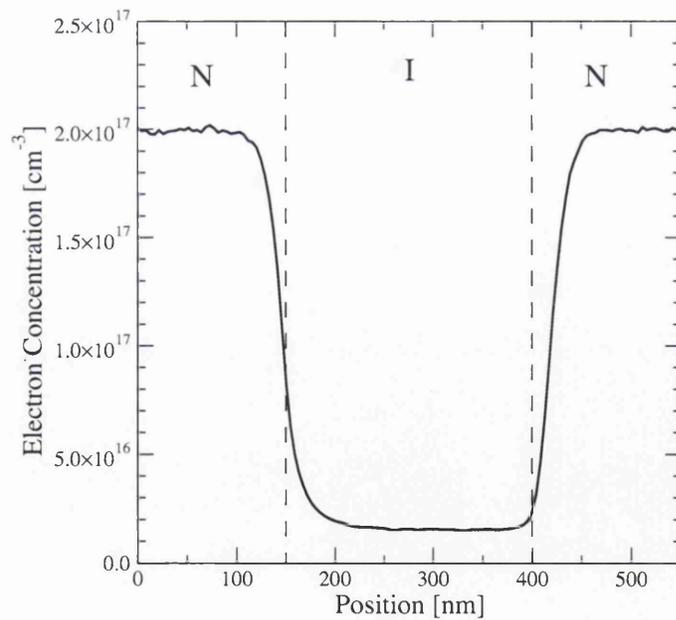


Figure 3.17: 1D profile of the electron concentration shown in figure 3.16 having averaged over the y and z directions. The boundaries between the N I N regions are marked by the dashed lines.

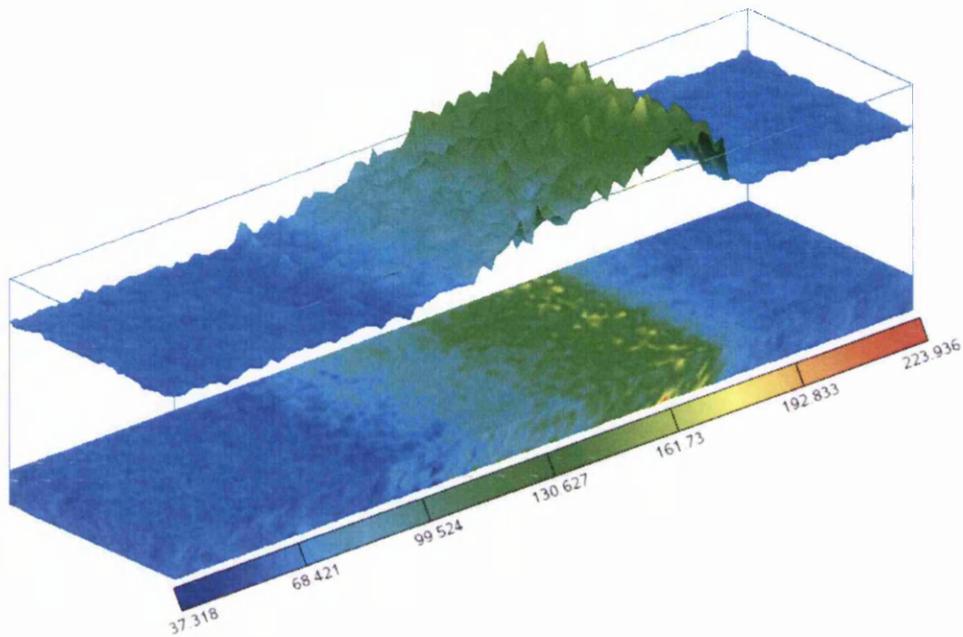


Figure 3.18: 3D simulated distribution of average electron energy in meV

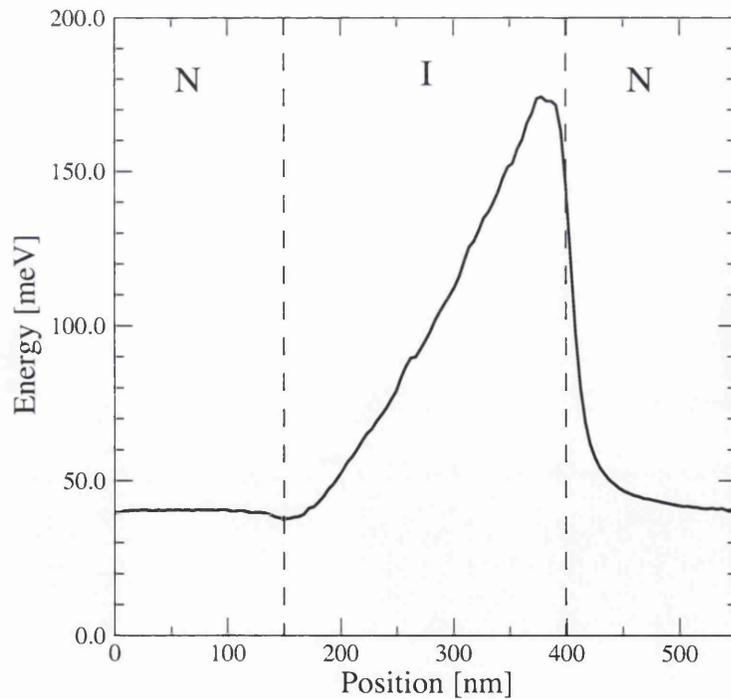


Figure 3.19: 1D profile of the electron energy distribution as shown in figure 3.18 having averaged over the y and z directions. The boundaries between the N I N regions are marked by the dashed lines.

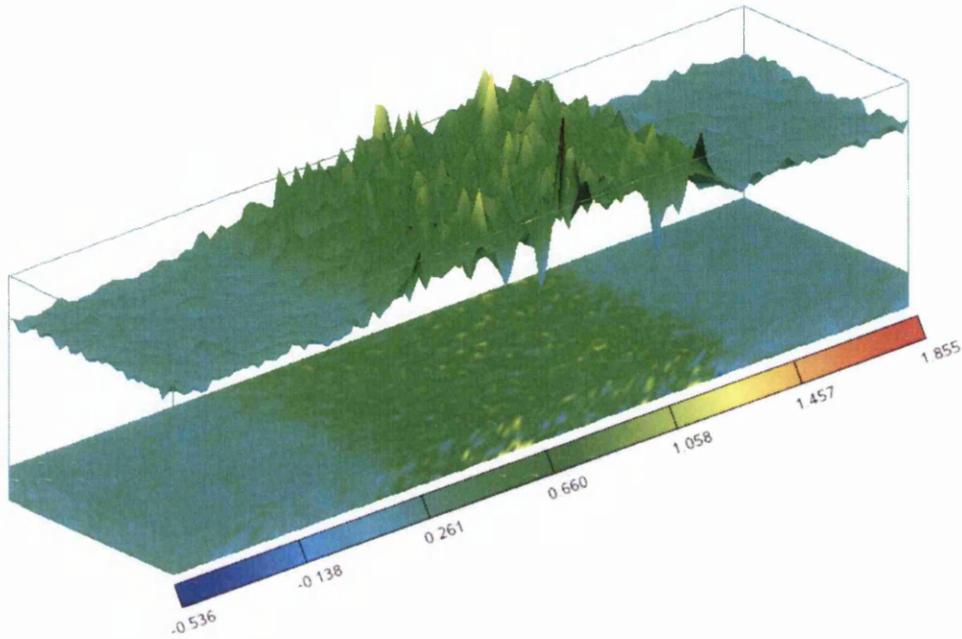


Figure 3.20: 3D simulated electron velocity distribution throughout the device. Units are in $\times 10^7 \text{ cm s}^{-1}$

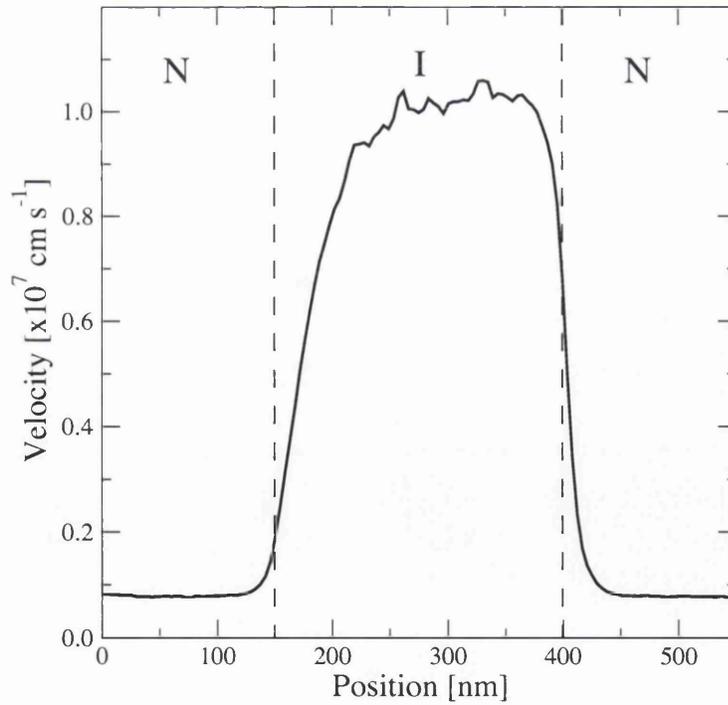


Figure 3.21: 1D profile of the electron velocity distribution as shown in figure 3.20 having averaged over the y and z directions. The boundaries between the N I N regions are marked by the dashed lines.

3.7 Conclusion

We have presented the basic models used in this 3D EMC simulator with discussion regarding their validity and propriety related to the problems addressed throughout this work. Validation of the phonon calibration through comparison of experimental bulk velocity-field curves, and a comparison with more sophisticated full band simulations given through bulk energy-field curves was presented. The main additions necessary to carry out device simulation were then discussed along with the nature of the discretization mesh, field calculation and field interpolation employed. Finally, an example simulation of an N-I-N diode with the 3D simulator was presented. The results are consistent and support the credibility of the developed simulator. In the next chapter the *ab initio* ionized impurity scattering method and application is verified before progressing to investigate the effects of discrete dopants in technologically important devices.

Band Independent Parameters	Symbol	Value
Transverse Sound Velocity	u_t	4.70 km s^{-1}
Longitudinal Sound Velocity	u_l	9.18 km s^{-1}
Mass Density	ρ	2.30 g cm^{-3}
Static Dielectric Constant	ϵ	11.9
Lattice Constant	a	5.43 \AA
Acoustic Deformation Potential	Ξ_{ac}	4.80 eV
X-Valley Parameters	Symbol	Value
Bandgap	E_g	1.12 eV
Nonparabolicity	α	0.43 eV^{-1}
Transverse Effective Mass	m_t^*	$0.191 m_e$
Longitudinal Effective Mass	m_l^*	$0.903 m_e$
Optical Coupling Constants		
<i>g1 process</i>	$(D_t K)_{g1}$	$0.450 \times 10^{10} \text{ eV m}^{-1}$
<i>g2 process</i>	$(D_t K)_{g2}$	$0.267 \times 10^{10} \text{ eV m}^{-1}$
<i>g3 process</i>	$(D_t K)_{g3}$	$10.75 \times 10^{10} \text{ eV m}^{-1}$
<i>f1 process</i>	$(D_t K)_{f1}$	$0.100 \times 10^{10} \text{ eV m}^{-1}$
<i>f2 process</i>	$(D_t K)_{f2}$	$1.950 \times 10^{10} \text{ eV m}^{-1}$
<i>f3 process</i>	$(D_t K)_{f3}$	$1.950 \times 10^{10} \text{ eV m}^{-1}$
Optical Phonon Energies		
<i>g1 process</i>	$\hbar\omega_{g1}$	$1.206 \times 10^{-2} \text{ eV}$
<i>g2 process</i>	$\hbar\omega_{g2}$	$1.853 \times 10^{-2} \text{ eV}$
<i>g3 process</i>	$\hbar\omega_{g3}$	$6.204 \times 10^{-2} \text{ eV}$
<i>f1 process</i>	$\hbar\omega_{f1}$	$1.896 \times 10^{-2} \text{ eV}$
<i>f2 process</i>	$\hbar\omega_{f2}$	$4.740 \times 10^{-2} \text{ eV}$
<i>f3 process</i>	$\hbar\omega_{f3}$	$5.903 \times 10^{-2} \text{ eV}$
Inter-valley Optical Coupling Constants		
<i>X - L</i>	$(D_t K)_{X-L}$	$2.34 \times 10^{10} \text{ eV m}^{-1}$
<i>X - Γ</i>	$(D_t K)_{X-\Gamma}$	$5.48 \times 10^{10} \text{ eV m}^{-1}$
Inter-valley Optical Phonon Energies		
<i>X - L</i>	$\hbar\omega_{X-L}$	$3.716 \times 10^{-2} \text{ eV}$
<i>X - Γ</i>	$\hbar\omega_{X-\Gamma}$	$2.189 \times 10^{-2} \text{ eV}$

Table 3.6: Band independent parameters and calibrated values for the minimum lying X-valley used in the EMC simulation. These values govern the transport processes effecting the majority of electrons in silicon and are used consistently throughout this work.

<i>L</i> -Valley Parameters	Symbol	Value
Bandgap	E_g	2.169 eV
Nonparabolicity	α	0.30 eV ⁻¹
Transverse Effective Mass	m_t^*	0.126 m_e
Longitudinal Effective Mass	m_l^*	1.634 m_e
Inter-valley Optical Coupling Constants		
<i>L</i> – <i>X</i>	$(D_t K)_{L-X}$	2.340×10^{10} eV m ⁻¹
<i>L</i> – <i>L</i>	$(D_t K)_{L-L}$	2.630×10^{10} eV m ⁻¹
<i>L</i> – Γ	$(D_t K)_{L-\Gamma}$	5.010×10^{10} eV m ⁻¹
Inter-valley Optical Phonon Energies		
<i>L</i> – <i>X</i>	$\hbar\omega_{L-X}$	3.716×10^{-2} eV
<i>L</i> – <i>L</i>	$\hbar\omega_{L-L}$	3.887×10^{-2} eV
<i>L</i> – Γ	$\hbar\omega_{L-\Gamma}$	2.090×10^{-2} eV
Γ -Valley Parameters	Symbol	Value
Bandgap	E_g	3.495 eV
Nonparabolicity	α	0.00 eV ⁻¹
Transverse Effective Mass	m_t^*	0.229 m_e
Longitudinal Effective Mass	m_l^*	1.987 m_e
Inter-valley Optical Coupling Constants		
Γ – <i>X</i>	$(D_t K)_{\Gamma-X}$	5.480×10^{10} eV m ⁻¹
Γ – <i>L</i>	$(D_t K)_{\Gamma-L}$	5.010×10^{10} eV m ⁻¹
Γ – Γ	$(D_t K)_{\Gamma-\Gamma}$	2.990×10^{10} eV m ⁻¹
Inter-valley Optical Phonon Energies		
Γ – <i>X</i>	$\hbar\omega_{\Gamma-X}$	2.189×10^{-2} eV
Γ – <i>L</i>	$\hbar\omega_{\Gamma-L}$	2.090×10^{-2} eV
Γ – Γ	$\hbar\omega_{\Gamma-\Gamma}$	2.568×10^{-2} eV

Table 3.8: Parameters for the *L* and Γ valleys. Though present within the simulation, little transport occurs in these valleys due to the large separation in energy from the *X*-valley. These values are provided for completeness.

Chapter 4

Ionized Impurity Scattering

4.1 Introduction

In the previous chapter, the bulk Monte Carlo (MC) method was presented. The adopted scattering models and their parameters were verified via comparison of experimental bulk silicon transport properties with simulation results. Discussion of the subsequent extension to 3D MC simulation was given and the developed 3D simulation code was verified through simulation of a simple N-I-N diode. In all above simulations, the ionized impurity scattering was introduced in a traditional way as a scattering mechanism dependent on the local, continuous, doping concentration and the local electron concentration. This method assumes the doping to be self averaging and treats scattering as an instantaneous two-body interaction. It can not accurately incorporate the effect of continuous many body scattering from a limited number of randomly positioned scattering centres within a small volume. In ultra-small MOSFETs, with channel lengths on the order of a few tens of nanometres, the discrete nature of the doping plays an important role and variation in both the position and the number of dopants within the channel leads to variation in device characteristics and parameters, such as threshold voltage V_T [41,42]. This adversely

effects the operation of circuits comprised of such unmatched devices. Predicting the magnitude of device parameter variation is therefore important in determining the functionality of circuits and systems containing an ensemble of fabricated devices. Simulations employing ionized impurity scattering rates can not resolve random dopant induced transport variations and an alternative method is therefore required.

In this chapter, such a suitable *ab initio* approach for introducing ionized impurity scattering is discussed and incorporated within the already developed 3D EMC device simulation tool. A review of the various approximations of ionized impurity scattering rates precedes the description of the *ab initio* model in section 4.2. Accurate integration of the equations of motion is particularly important with respect to properly resolving *ab initio* Coulomb scattering from point charges and is discussed in section 4.3 and validated in section 4.4. The model itself is discussed in detail in section 4.5. Finally, the whole method is validated via the *ab initio* reproduction of the low field bulk concentration dependent mobility in section 4.6, and via comparison of results obtained from the simulation of an 'atomistic' N-I-N diode with the 'continuous' diode simulated in chapter 3 in section 4.7.

4.2 Ionized Impurity Scattering Rates

The scattering rates for the mechanisms described in chapter 3 are all constructed from 1st order perturbation theory following the Born approximation. This is based upon the assumptions that a scattering event is attributable to a small perturbation in the periodic crystal potential, that such an event occurs over a short enough time as to be considered instantaneous, and that individual events are infrequent in time and so can be treated independently. Within this framework, neglecting external fields, scattering instantaneously results in the transition of an electron from one time independent Bloch state to another.

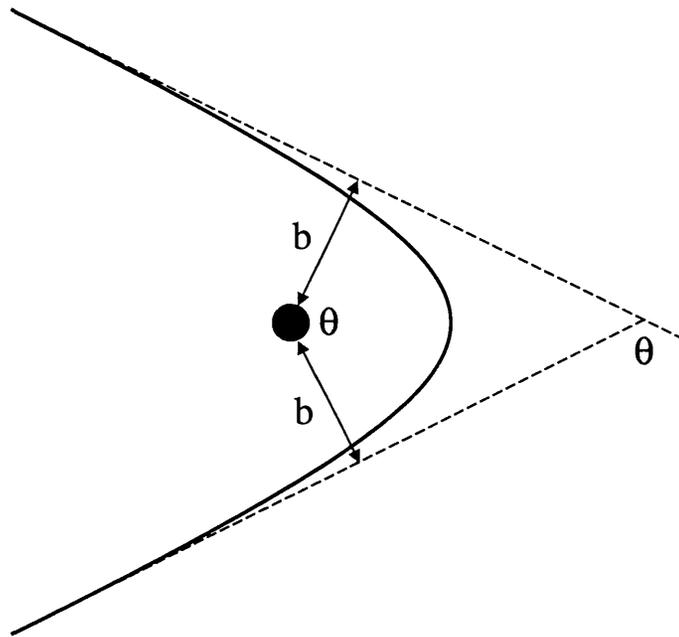


Figure 4.1: Definition of the impact parameter b and the scattering angle θ in the Rutherford model of scattering. Electrons propagate in a hyperbolic orbit in the field of a single point charge, here taken to be positive.

For the phonon mechanisms considered, the Born approximation is a good one. However, the interaction with a pure Coulomb potential, which is the perturbation potential for ionized impurity scattering, is infinite in range and results in an infinite interaction cross section. This implies a scattering event that must be considered extended in time. Additionally at high doping concentrations, significant scattering from multiple impurities means that treating interactions independently is also inaccurate. Though there exist solutions to the scattering rate due to an isolated ionized impurity in the form of phase-shift analysis, which is not limited by the Born approximation [127], the simplest and most common forms of ionized impurity scattering models applied in MC simulations are, however, based upon this approximation.

Although the ionized impurity scattering rate is calculated quantum mechanically, a classical description clarifies the choice of scattering potential and the scattering outcome.

Let us consider one electron in the electrostatic potential of a singly ionized impurity atom, approximated at a distance by the Coulomb potential. The impact parameter b can be defined as the perpendicular distance from the scattering centre to the line asymptotic to the initial, or final, trajectory with the ion. This is illustrated in figure 4.1. A larger impact parameter results in a smaller angle of deflection due to the reduced interaction between the charges. In the Rutherford theory of Coulomb scattering, the angle of deflection θ is determined completely by the impact parameter through the relation

$$\theta = 2 \arctan \left(\frac{e^2}{8\pi\epsilon_r\epsilon_0 b \epsilon(\mathbf{k})} \right) \quad (4.1)$$

where $\epsilon(\mathbf{k})$ is the incident electron energy, e the electronic charge, ϵ_0 the permittivity of free space and ϵ_r the relative permittivity of the surrounding medium. The infinite collisional cross section favours small angle deflections, hence small momentum transfer collisions, due to the greater probability of interactions with a large impact parameter. In this way a reconciliation with the Born approximation is customarily made by approximating the interaction with the impurity ion potential as being comprised of two regimes. The first is a weak long-range interaction leaving the electron state unaltered, and the second is a strong, infrequent, short-range two-body interaction which scatters the electron into a new state [128]. This interpretation removes the problematic infinite scattering cross-section by neglecting small angle deflections, and is consistent with the Born approximation in that the interaction can be approximated as an instantaneous two-body interaction, so long as the scattering potential falls off rapidly at large distances and the electron energy is large [129]. Suitable representations for the short range interaction potential then need to be used within the perturbation theory. The representations commonly used in MC simulation lead to the Brooks-Herring, Conwell-Weisskopf and Third Body Exclusion ionized impurity scattering models which are discussed next.

4.2.1 Brooks-Herring Model

The Brooks-Herring (BH) model relies on the fact that the electrostatic potential associated with a bare Coulomb charge will be screened by surrounding mobile charges, if present, to limit the range of the interaction. Screening partly arises from the polarization of the local material and is accounted for through the associated dielectric constant. This only acts to reduce the magnitude of the bare Coulomb potential without effecting the range of the interaction. More important is the screening by mobile charges. In the simplest case this is treated as static screening and results in an exponential decay of the potential over a characteristic screening length λ , such that the potential is given by

$$V = \frac{q}{4\pi\epsilon r} \exp(-\beta r) \quad (4.2)$$

where β is the inverse screening length, simply related to λ by

$$\beta = \frac{1}{\lambda} \quad (4.3)$$

The characteristic screening length in its most basic form is assumed to be the Debye length

$$\lambda_D = \sqrt{\frac{\epsilon k_B T_{el}}{ne^2}} \quad (4.4)$$

where n is the local electron density, T_{el} is the local electron temperature and the remaining symbols have their usual meaning. The screened Coulomb potential is plotted alongside the bare Coulomb potential in figure 4.2, for an electron concentration of $1 \times 10^{18} \text{ cm}^{-3}$.

This choice of screening length does not account for degeneracy and at higher carrier

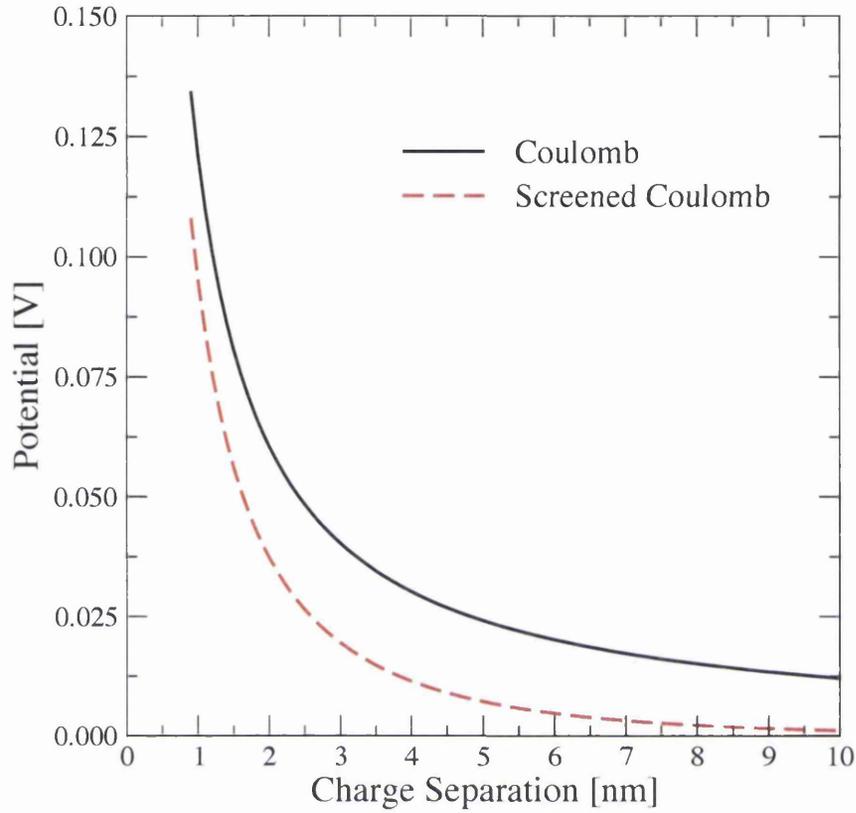


Figure 4.2: Comparison of the screened Coulomb potential used in the BH model of impurity scattering to that of the unscreened Coulomb potential. Screening by mobile charges causes the potential to drop off faster and reduces the range of the interaction.

concentrations the more general form for β should be used, where

$$\beta^2 = \frac{ne^2}{\epsilon k_B T_{el}} \frac{F_{-1/2}(\eta)}{F_{1/2}(\eta)} \quad (4.5)$$

Here F_i is the Fermi-Dirac integral of order i , given by

$$F_i(\eta) = \int_0^\infty \frac{x^i}{1 + \exp(x - \eta)} dx \quad (4.6)$$

and η is the reduced Fermi energy, given by

$$\eta = \frac{E_c - E_f}{k_B T_{el}} \quad (4.7)$$

where E_c and E_f are the conduction band energy and Fermi energy respectively. The degeneracy according to equation 4.5 increases the screening length, and must be included in order to satisfactorily describe mobility at high doping concentrations [130].

Additionally, the assumption of static screening is only valid in cases where the screening charges can respond faster than the perturbation in the potential as a charge to be scattered passes through the scattering centre. This response time is characterised by the dielectric relaxation time and is proportional to the mobile charge density. At low temperatures, 80K and below, the dynamic screening is important but in general can be neglected at room temperatures [131].

Using the screened potential, the scattering rate for a single impurity centre can be determined within the Born approximation. To obtain the total scattering rate from multiple impurities this rate is simply multiplied by the concentration of the scattering centres. This is a major drawback of the BH model in that it can not take into account simultaneous scattering from multiple impurities. In this way the scattering is unrealistically treated at high concentrations where scattering from multiple impurity ions often occurs [132].

4.2.2 Conwell-Weisskopf Model

Opposite to the BH approach, the Conwell-Weisskopf (CW) model neglects screening but limits the range of the interaction by the proximity of the neighbouring ionized impurities. As an electron moves further away from one impurity, it approaches another. The CW approach assumes that an electron only interacts with one scattering centre at a time - the closest one. For a given concentration N_I of ionized impurities, the average separation

is $N_I^{-\frac{1}{3}}$, assuming the ionized impurities are uniformly distributed in the centre of a cube of length $N_I^{-\frac{1}{3}}$. An electron therefore only interacts with the Coulomb potential of an impurity up to separations of $\frac{1}{2}N_I^{-\frac{1}{3}}$ before it is considered to be participating in another two-body interaction. Consequently, the interaction potential is then given by

$$V = \frac{q}{4\pi\epsilon r} \quad 0 < r \leq \frac{1}{2}N_I^{-\frac{1}{3}} \quad (4.8)$$

This potential is plotted alongside the screened potential of the BH model for a doping concentration of $1 \times 10^{18} \text{ cm}^{-3}$ in figure 4.3. This limits the impact parameter to $b_{max} = \frac{1}{2}N_I^{-\frac{1}{3}}$, consequently defining a minimum angle of deflection. In this way the infinite collisional cross section is removed, though somewhat artificially.

The CW potential better represents scattering in highly doped regions where the separation between ionized impurities is small and where a two body interaction between an electron and an alternate impurity ion is likely to dominate for large impact parameters. Similarly to the BH model, having obtained the scattering rate for a single impurity within the CW approximation, the total scattering rate is found by multiplying by the local impurity concentration. This again suffers from neglecting possible scattering from multiple centres. Additionally, screening from mobile charges is also disregarded but should be included to properly account for the scattering potential in regions of high mobile charge concentration within a device.

4.2.3 Third Body Exclusion Model

The merits of the BH and CW approaches are unified in the statistical screening, or third body exclusion (TBE), model. In this case the impurity potential is taken as the screened Coulomb potential of the BH model. This is connected to the CW model by incorporating the probability that no other scattering centre exists that is closer to the electron than

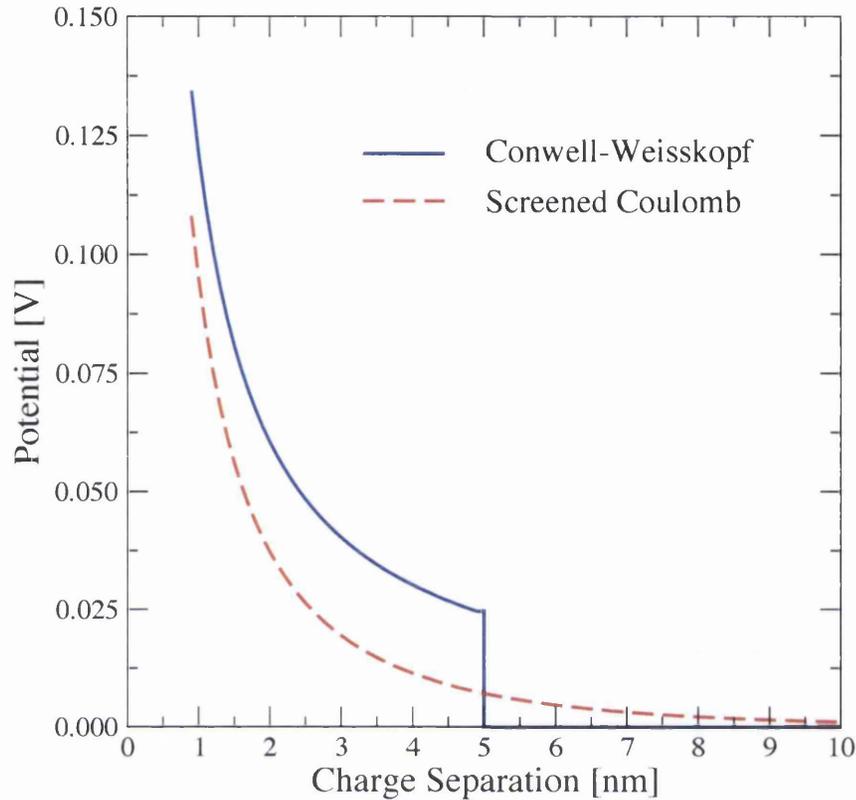


Figure 4.3: Comparison of the interaction potential used in the Conwell Weisskopf model with that of the Screened Coulomb potential as plotted in fig 4.2. At a doping concentration of $1 \times 10^{18} \text{ cm}^{-3}$ the average separation between dopants is 10 nm , therefore the range of interaction of a single dopant is 5 nm .

the scattering centre currently under consideration [133]. The probability that no other scattering centre exists with an impact parameter in the range b and $b + db$ is given by

$$p = 1 - 2\pi N_I a b db \quad (4.9)$$

where N_I is the concentration of impurity ions with average separation a , where a is given by

$$a \approx (2\pi N_I)^{-\frac{1}{3}} \quad (4.10)$$

Equation 4.10 is an approximation to the mean separation of impurities based upon the probability of finding within a spherical volume of radius b centred about one impurity, a secondary impurity with an impact parameter less than b [128]. This differs from the average separation used in the CW model, $b_{max} = \frac{1}{2}N_I^{-\frac{1}{3}}$, assuming a uniform distribution of impurities associated each with a cubic volume. The difference between the models is however less than 10%. If $P(b)$ is the probability that no other scattering centre exists with impact parameter less than b and, consequently, $P(b + db)$ the probability that none exist with impact parameter less than $b + db$, then

$$P(b + db) = P(b) p \quad (4.11)$$

and it follows that

$$P(b) = \exp(-\pi N_I a b^2) \quad (4.12)$$

This probability excludes the presence of a closer secondary scattering centre as the impact parameter increases. As such it incorporates screening with the probability of a nearest neighbour scattering centre. In the limit of low concentration of scattering centres, this result reproduces the BH result, while at high concentrations it reproduces the CW result. As such it merely extends the range of the two body interaction approximation and does not account for scattering from multiple scattering centres.

4.2.4 Limitation of scattering rates

The scattering rates previously described are all based upon a two-body interaction with varying approximations determining the range of the interaction and the form of the scattering potential. This isolated two body process is extended to describe the total rate of such two-body scattering events for a distribution of ionized impurities by simply mul-

tipling the result by the local concentration of impurities. This neglects the cumulative influence of all impurities other than the nearest one. However, local inhomogeneities and clustering in the doping distribution over small volumes often violates such assumptions. At high doping concentrations, simultaneously scattering from multiple ionized impurities then becomes a significant unaccounted effect in the described models. Additionally, the actual deterministic scattering resulting from a distribution of exactly located ionized impurities within a device is here approximated by a stochastic process based upon the average impurity concentration. In this way scattering rates are unable to reproduce transport variation in devices due to variation in dopant positions.

In order to fully evaluate the impact of random dopants, the position dependent ionized impurity scattering must be treated in a deterministic way. This requires a full 3D MC device simulation in order to capture both the electrostatic and transport effects of unique 3D dopant configurations from device to device. This is achieved through an *ab initio* approach in which the traditional scattering rates are removed, and instead ionized impurity scattering is treated through the real space trajectories of carriers in the potential landscape associated with the unique arrangement of charges.

Within this model, scattering from multiple impurities is automatically accounted for in the classical picture of an electron orbit. The dynamic screening is also automatically taken into account, removing the necessity of making prior assumptions about the distribution function as used in deriving the screening length for use in scattering rates [75]. The *ab initio* Coulomb scattering method as applied to electron-dopant interactions can similarly be applied to electron-electron interactions, automatically accounting for electron-electron scattering within simulations [77].

Since results from the *ab initio* Coulomb scattering approach rely upon the real space trajectories of electrons, the integration of the equations of motion must be able to resolve the interactions with sufficient accuracy. This is addressed next.

4.3 Integrating the Equations of Motion

The phase space coordinates, \mathbf{r} and \mathbf{k} , of each electron are repeatedly advanced through discrete intervals in time, Δt , for the duration of the simulation. The time step should be long enough for efficient simulation but small enough as to maintain accuracy and stability of the numerical integration of the carrier trajectories. Stability is normally enforced in self-consistent simulations by ensuring that the maximum plasma frequency expected in a simulation is adequately sampled (see chapter 3). Within the *ab initio* EMC we must also ensure that the time step is small enough and the integration routine accurate enough to properly resolve the particle trajectories in the high electric fields, and large rate of change of fields, associated with the sharply resolved Coulomb potential close to the discrete charges.

Accurately integrating the equations of motion for an ensemble of interacting particles has been a problem studied in plasma and gravitational simulations for many years, though in many cases short range interaction are deliberately omitted in cases when only the long range interactions play significant roles [134]. A variety of numerical methods and corresponding routines were tested in this work, ranging from very basic to high order. Their suitability and accuracy was tested in simple cases representative of the requirements during device simulation. These include reflection from a boundary perpendicular to a large electric field and propagation in the field of a fixed point charge. A basic description of various numerical integration methods is given in the following sections, detailing their limitations within the simulation framework.

4.3.1 Euler Integration

Introducing the notation adopted here, the position and wave vector of an electron at time t are written as \mathbf{r}_t and \mathbf{k}_t , respectively. These coordinates determine the force $\mathbf{F}(\mathbf{r}_t)$

and velocity $\mathbf{v}(\mathbf{k}_t)$ at time t respectively, used for propagating the coordinates to time $t + \Delta t$. The simplest form of integration is the Euler method described by the following relationships

$$\mathbf{r}_{t+\Delta t} = \mathbf{r}_t + \mathbf{v}(\mathbf{k}_t) \Delta t \quad (4.13)$$

$$\mathbf{k}_{t+\Delta t} = \mathbf{k}_t + \mathbf{F}(\mathbf{r}_t) \Delta t \quad (4.14)$$

This method uses a single evaluation of the net force and the velocity at time t in order to advance them to $t + \Delta t$, but is only accurate to order Δt . This requires too small a time step for achieving a reasonable accuracy when other methods can tolerate larger time steps and achieve the same accuracy.

4.3.2 Runge-Kutta Schemes

Better accuracy in integrating the equations of motion can be achieved by using higher order methods. Two such commonly used higher order schemes are the 2nd and 4th order Runge-Kutta methods [135]. The 2nd order scheme may be written as

$$\mathbf{k}_{t+\Delta t} = \mathbf{k}_t + \mathbf{F}(\mathbf{r}') \Delta t \quad (4.15)$$

$$\mathbf{r}_{t+\Delta t} = \mathbf{r}_t + \mathbf{v}(\mathbf{k}') \Delta t \quad (4.16)$$

with

$$\mathbf{r}' = \mathbf{r}_t + \mathbf{v}(\mathbf{k}_t) \frac{\Delta t}{2} \quad (4.17)$$

$$\mathbf{k}' = \mathbf{k}_t + \mathbf{F}(\mathbf{r}_t) \frac{\Delta t}{2} \quad (4.18)$$

This method is also referred to as the midpoint method because the time derivatives are evaluated not only at the beginning, but half way through the time interval. Higher order Runge-Kutta methods are very common in numerical integration and extend this process

to take weighted averages of multiple time derivatives estimated between t and $t + \Delta t$.

However, in the case of 3D self consistent EMC simulations, the evaluation of the force at any point other than \mathbf{r}_t can lead to unphysical results. This is as a result of too few, or even a single, particle per mesh cell moving in their own mesh resolved potential. Evaluation of the field at a point other than the current position of the carrier within the solution of the Poisson equation leads to a mesh dependent force that accelerates particles. Effectively electrons un-physically 'feel' a force imparted by a copy of themselves. The addition of more simulated particles smoothes the potential and reduces this error. In the *ab initio* approach, the accurate treatment of electron-electron scattering demands that one particle represents one electron. In such situations we are unable to choose and increase the number of particles used in the simulations. For these reasons, integration schemes are limited to those that sample the field once per solution of Poisson's equation at the position of the carriers corresponding to the solution.

4.3.3 Classical Dynamics

One integration scheme that relies on single force evaluation is the familiar form of Newton's equations of motion

$$\mathbf{r}_{t+\Delta t} = \mathbf{r}_t + \mathbf{v}(\mathbf{k}_t) \Delta t + \frac{\mathbf{F}(\mathbf{r}_t)}{2m} \Delta t^2 \quad (4.19)$$

$$\mathbf{k}_{t+\Delta t} = \mathbf{k}_t + \mathbf{F}(\mathbf{r}_t) \Delta t \quad (4.20)$$

This approach includes the impact of acceleration during the propagation step on the final position and has been used in similar molecular dynamics EMC routines to study electron-electron interactions in bulk semiconductors [75]. It, however, assumes that the mass is constant during propagation which within the semiclassical equations of motion

is in general not true, especially at high fields.

For the analytic band model employed, the dispersion relation around the minima is defined by the non-parabolicity factor α . In this approximation the term representing mass, defined by the ratio of crystal momentum to velocity, is given by

$$m = m^* (1 + 2\alpha\varepsilon(\mathbf{k})) \quad (4.21)$$

where m^* is the effective mass at the band minimum and $\varepsilon(\mathbf{k})$ the dispersion relation (section 3.3.1). For a non parabolic band with $\alpha > 0$, the mass increases with energy. In large fields, such as the field perpendicular to the gate in the channel of a MOSFET in accumulation, the change in energy, and hence effective mass, during one time step is large. Overlooking this is at best inaccurate, but in the worst case can lead to a loss of energy conservation over a closed path. As an example, over the course of reflection from the insulator interface in the channel of a MOSFET, the mass is underestimated during the acceleration in the field towards the boundary and, after subsequent reflection, overestimated following the same path back. This leads to violation of the energy conservation in the channel through inaccurate evaluation of the position and subsequent unphysical loss of carrier energy. This failure is due to the lack of time symmetry in the propagation and therefore any integration scheme used must also resolve this problem

4.3.4 Leap-Frog

The most commonly used particle propagation routine in EMC simulations, due to both its simplicity and accuracy, is the leap-frog method [134]. This method is applied to plasma simulations, molecular dynamics simulations and gravitational simulations all with great success. It utilises only one force evaluation and is time centred. The discretized time centred equations of motion remove the problem associated with reflection from the interface

and loss of energy discussed above. Position and wave vector coordinates are stored at staggered intervals in time, separated by $\Delta t/2$. Each is propagated forward in time using the following simple relations

$$\mathbf{k}_{t+\Delta t/2} = \mathbf{k}_{t-\Delta t/2} + \mathbf{F}(\mathbf{r}_t) \Delta t \quad (4.22)$$

$$\mathbf{r}_{t+\Delta t} = \mathbf{r}_t + \mathbf{v}(\mathbf{k}_{t+\Delta t/2}) \Delta t \quad (4.23)$$

Leap-frog integration has the drawback that the coordinates are not synchronised. This is not necessarily a problem if only the particle trajectory is required or, in the case of MC simulation, if the difference in the wave vector $\mathbf{k}_{t+\Delta t/2}$ over half a time step is not enough to significantly effect the scattering rate at the point $\mathbf{r}_{t+\Delta t}$. The technique is commonly used in collisionless systems where the propagation iterates indefinitely. This is not the case in semiconductor simulation where the process is broken by scattering events and interactions with boundaries in a device simulation. In this case it becomes simpler to consider the wave vector and position at the same point in time and propagate to a collision event, this is achieved in the Velocity-Verlet algorithm [136].

4.3.5 Velocity-Verlet

The Velocity-Verlet algorithm for propagation can be reduced to the leap-frog algorithm. Whereas the leap-frog algorithm propagates both the position and momentum in one step, storing results at staggered time intervals, the Velocity-Verlet algorithm propagates the position in one step while the momentum in two steps, such that the results are synchronous. Propagation of position and momentum follows a three step process. First the momentum at time t is advanced by $\Delta t/2$ using the field at time t . Propagation of position then follows identically with leap-frog, using the velocity value at time $t + \Delta t/2$, and

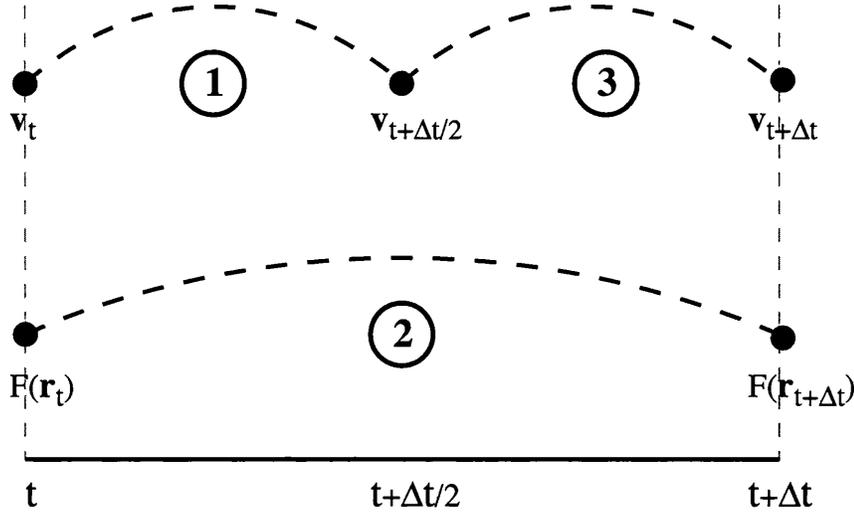


Figure 4.4: Velocity-Verlet algorithm propagates momentum for (1) half the time step before (2) a single propagation of position using the velocity at $t + \Delta t/2$ and finally (3) updates the momentum using the field at $t + \Delta t$.

finally the momentum is updated to the same point in time as the position by using the electric field evaluated at the new position. This process is illustrated in figure 4.4 and is expressed by the following relationships

$$\mathbf{k}_{t+\Delta t/2} = \mathbf{k}_t + \mathbf{F}(\mathbf{r}_t) \frac{\Delta t}{2} \quad (4.24)$$

$$\mathbf{r}_{t+\Delta t} = \mathbf{r}_t + \mathbf{v}(\mathbf{k}_{t+\Delta t/2}) \Delta t \quad (4.25)$$

$$\mathbf{k}_{t+\Delta t} = \mathbf{k}_{t+\Delta t/2} + \mathbf{F}(\mathbf{r}_{t+\Delta t}) \frac{\Delta t}{2} \quad (4.26)$$

Like the leap-frog algorithm, Velocity-Verlet is time centred, utilises only one force evaluation per solution of Poisson's equation but, beneficially, its coordinates are synchronised.

Scattering events and interactions with physical boundaries break the propagation cycle. Since information regarding the new field can not be obtained until the result of the scattering event is obtained, propagation to the scattering event during a time step can

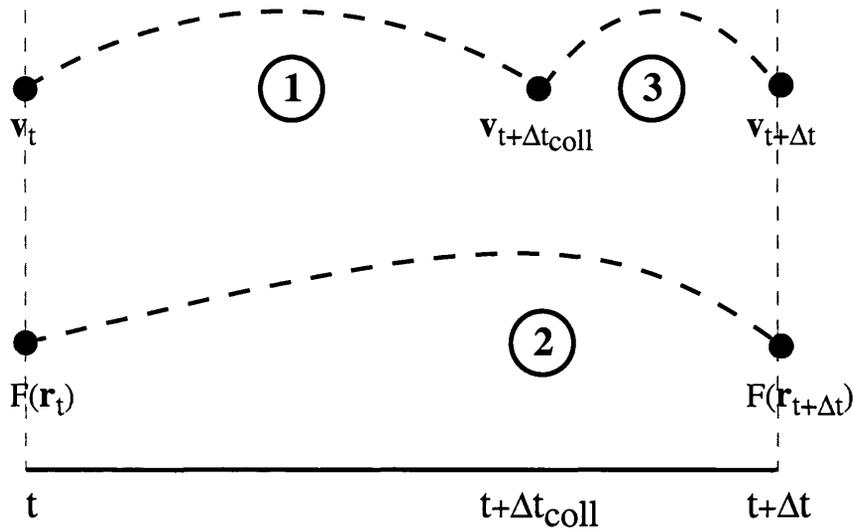


Figure 4.5: During scattering, propagation proceeds to the scattering event, time $t + \Delta t_{coll}$, before the new position can be found and the momentum updated. This breaks the time symmetry of the propagation routine.

only make use of the initial field. The loss of time centring in this case is unimportant as we do not require to propagate back and recover the state before scattering, though some error will be introduced by this deviation from the propagation cycle. The error in the values after propagation resulting from the departure from the iterative process to accommodate scattering and physical boundaries should be minimal for small time steps. The effect of scattering events on the time integration cycle is shown in figure 4.5.

Extension to this scheme can be made with the Beeman algorithm [136]. This in turn uses weighted values of the field from the previous time step, as well as the current field and the field at the subsequent time step. This should give better accuracy in energy conservation but requires the storing of multiple electric field values for each electron and is cumbersome to incorporate with carrier scattering. Should greater accuracy be necessary, this integration scheme may be considered.

2D EMC simulations allow larger numbers of particles per cell and yields a smoother

potential, justifying the use of higher order Runge-Kutta schemes. However the Velocity-Verlet algorithm best satisfies the requirements for propagation in the self consistent 3D *ab initio* EMC simulations. It is second order accurate and time centred while only requiring one field evaluation per solution of Poisson's equation. It remains to verify the use of this approach provides enough accuracy.

4.4 Validation of Propagation

Inaccurate integration of the equations of motion leads to physically inaccurate simulations and undermines the confidence in the results. Whereas collisionless simulations, like typical plasma simulations, require the accurate calculation of particle trajectories throughout the whole simulation, MC simulation of semiconductor transport only requires accurate trajectories between successive collisions. This makes the task of integrating the equations of motion somewhat easier. However, the large rate of change in the electric field associated with discrete charges within the *ab initio* EMC method introduces additional constraints as compared with the traditional MC method. It is therefore important to test and trust the integration method.

The integration methods previously discussed are applied to a series of simple test cases addressing common problems within the *ab initio* EMC simulation. These are used to highlight the aforementioned shortcomings of the discussed methods. The results verify that a lack of time centring in the propagation combined with an energy dependent mass results in loss of energy upon reflection, validating the choice of the Velocity-Verlet method. It is also shown that, while lower order methods poorly conserve energy and angular momentum in an orbit, the Velocity-Verlet approach is comparable in accuracy to higher order Runge-Kutta schemes when tested in an analytical Coulomb field. Scattering in an analytical Coulomb field is then considered for the Velocity-Verlet algorithm alone

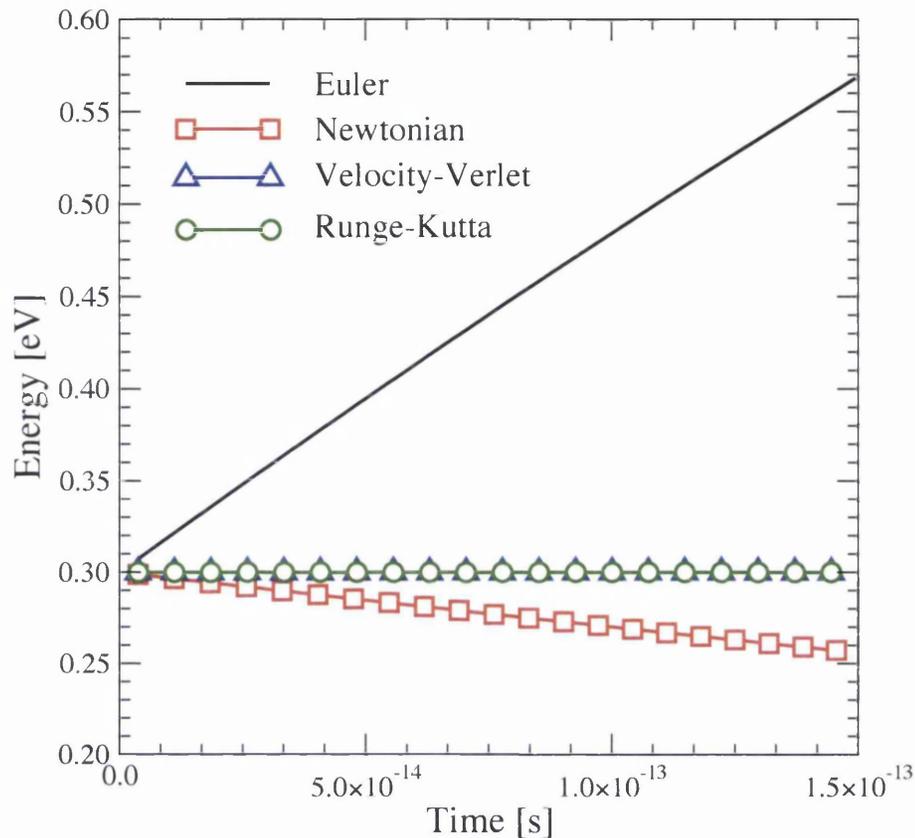


Figure 4.6: Total energy as a function of time for the simulated reflection off a smooth surface in a large electric field using different integration schemes and a non-parabolic band. Only the Runge-Kutta schemes and Velocity-Verlet maintain conservation of energy.

as this is of primary importance to further simulation.

4.4.1 Reflection at surface

In order to accurately conserve energy if a reflection occurs during the integration time step Δt , electrons must be propagated from the start time t to the moment they hit the boundary. The momentum is then reflected normal to the boundary and propagation continues to time $t + \Delta t$. This is in contrast to the simplistic method where propagation for the full time results in a final position beyond the boundary, followed by the reflection of

the particle's coordinates about the boundary.

The method is applied without loss of generality to a 1D test case considering an electron driven in a constant field towards a purely reflective boundary. Such a test is representative of the dynamics under the gate in later MOSFET simulations where surface roughness scattering is not included. The total energy as a function of time for different integration methods using a non-parabolic band are plotted in figure 4.6. As stated before, for non-parabolic bands energy conservation is only ensured with a properly time-centred integration routine. The lack of time centring results in a rapid loss of energy in the Newtonian dynamics scheme, which treats parabolic bands exactly, while the Velocity-Verlet and Runge-Kutta schemes perform well. In both parabolic and non-parabolic cases, the Euler scheme is too inaccurate and results in an unphysical gain in energy.

4.4.2 Propagation in an orbit

Accuracy in treating the interactions of point charges is paramount to valid simulation in this work. The propagation routines discussed and tested in the previous sections are again employed to propagate an electron in an elliptical orbit about a positive point charge assuming a parabolic band. An elliptical orbit is chosen, being more problematic than propagation in a circular orbit. The eccentricity of the orbit is 0.5 while the separation at closest approach is $2nm$. An analytical Coulomb field is used in order to eliminate interpolation errors from a mesh based solution and to allow a direct comparison of the techniques. The total energy throughout the propagation for the four different integration schemes is shown in figure 4.7. The simpler Euler and Newtonian dynamics schemes are unable to accurately treat the rapidly varying field close to the point charge resulting in a significant gain in energy during the closest approach of each orbit. This, however, is not a significant problem in either the Velocity-Verlet or 4th order Runge-Kutta schemes

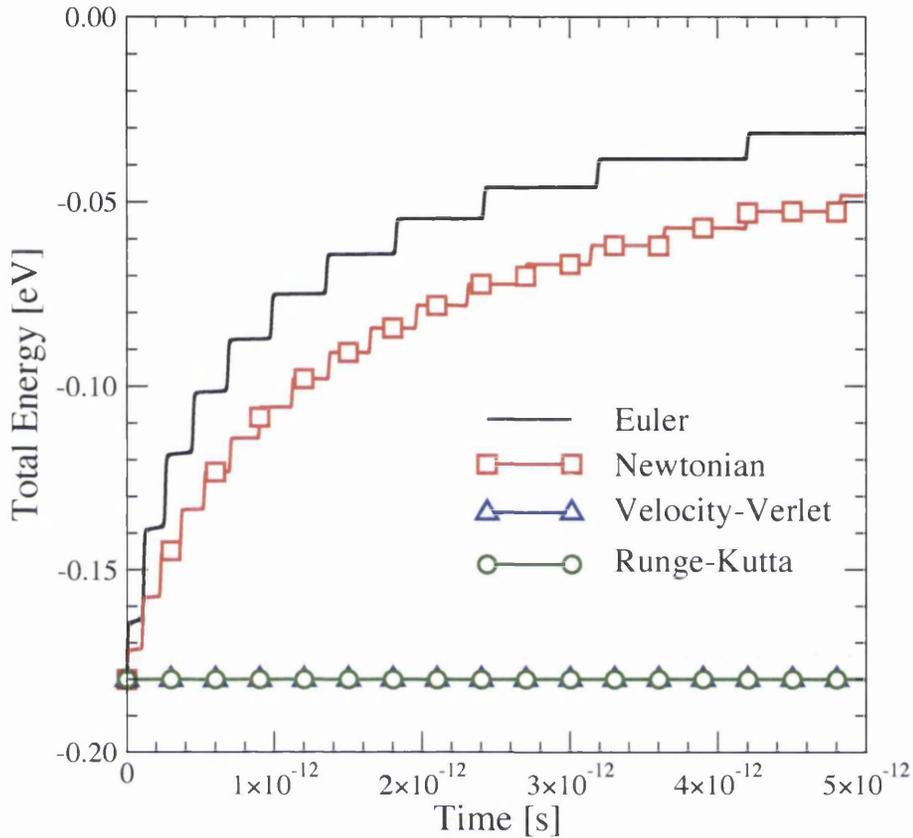


Figure 4.7: Total energy during propagation in an elliptical orbit using various integration schemes. Again Euler and Newtonian dynamics schemes perform poorly. Both the Velocity-Verlet and 4th order Runge-Kutta maintain conservation of energy.

which both preserve the total energy well.

4.4.3 Coulomb scattering

Following the classical picture of electron scattering illustrated in figure 4.1, the reproduction of the scattering angle dependence (equation 4.1) upon impact parameter via simulation using the Velocity-Verlet integration scheme is investigated. Reproduction of the scattering angle dependence is essential if ionized impurity scattering is to be accurately captured within *ab initio* EMC simulation. Both electron scattering from a fixed singly

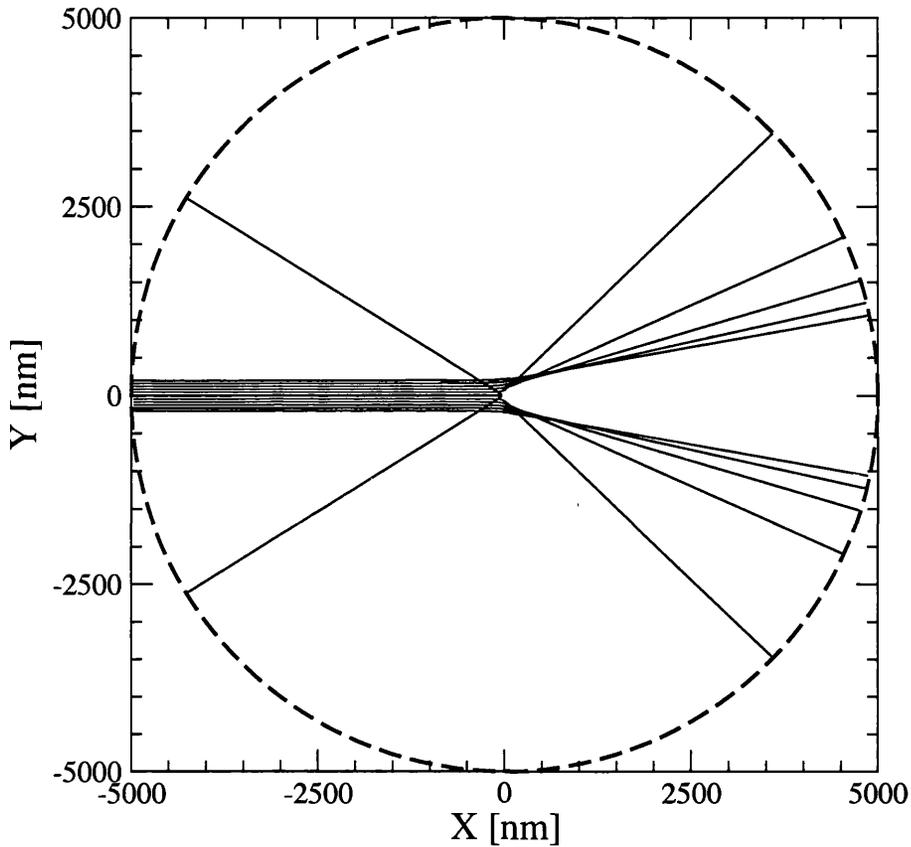


Figure 4.8: Simulated trajectories of electrons in the field of a central negative point charge. Electrons are injected with parallel velocities from the left and the dashed line marks the extent of the simulation domain.

negative or positive ion are considered, representing the interactions with an acceptor or donor respectively in later device simulation.

Electrons are initialized over a range of impact parameters at the edge of a circular simulation domain with parallel velocities determined by the incident total energy. The boundary of the simulation domain is sufficiently far from the centrally located point charge such that the initial and final velocities approximate the asymptotic velocities 'before' and 'after' the collision. The radius of the simulation domain is $5\mu\text{m}$ while the range in impact parameters considered is 200nm as shown in figure 4.8. Each electron

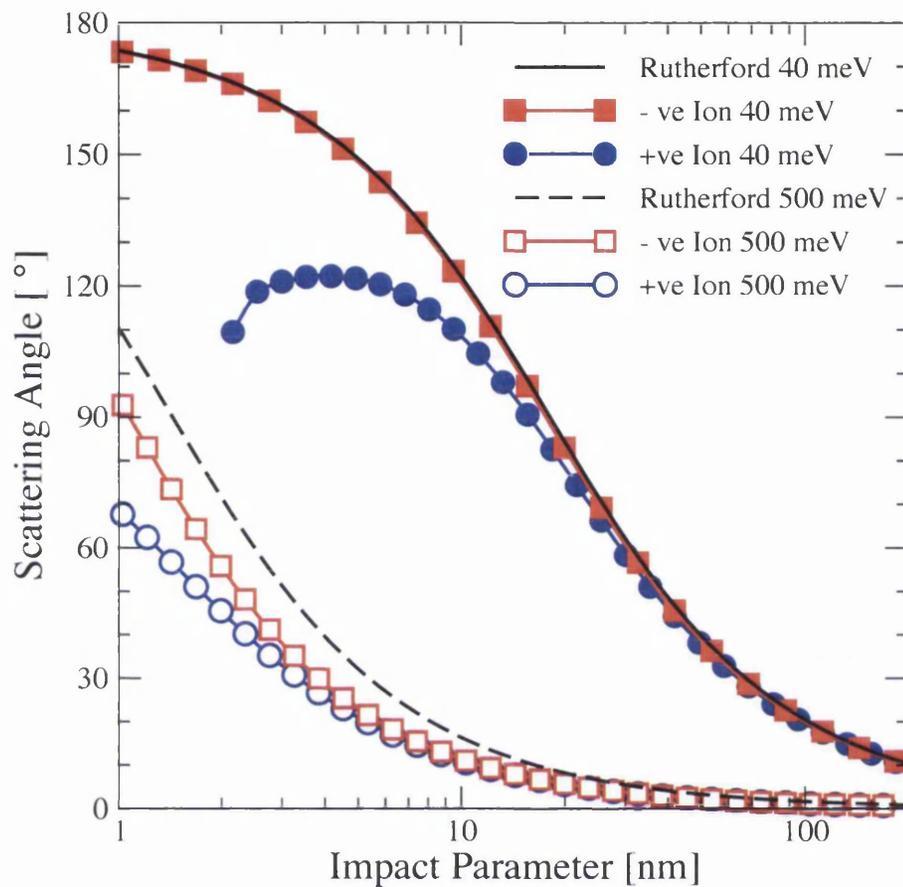


Figure 4.9: Variation of simulated scattering angle versus impact parameter compared with the Rutherford result. Results for both a positive and negative central charge are shown for two incident electron energies of 40 meV and 500 meV.

is propagated in time steps of 0.1 fs in the analytic Coulomb potential associated with the point charge until it leaves the simulation domain, whereupon the corresponding angle of deflection and percentage change in kinetic energy is determined. The simulated scattering angle dependence upon impact parameter for both a positive and a negative ion and considering incident electron energies of 40 meV and 500 meV are compared with the Rutherford formula (equation 4.1) in figure 4.9.

Equation 4.1 relating the scattering angle to the impact parameter and incident electron energy is independent of the sign of the charge on the ion. Within the simulation,

differences due to inaccurate integration become apparent. At large impact parameters for both incident energies, the simulated scattering from positive and negative ions agree with equation 4.1 due to the slowly varying fields and resultant accuracy in the integration of the electron's equations of motion. At low incident energy, the agreement is maintained over the entire range of impact parameters for the negative charge due to the repulsive nature of the interaction limiting the propagation in the rapidly varying field close to the scattering centre. At small impact parameters, the attractive nature of the positive ion results in propagation through the rapidly varying field near the ion, leading to inaccurate integration of position and energy. The simulated scattering angle for the case of the positive charge and low incident energy is not shown for impact parameters below 1 nm as the error in propagation becomes too significant. This error is reduced at higher incident energies due to the reduced propagation time in the high field region. For the negative ion, higher incident energy allows for penetration into the rapidly varying field and inaccurate integration is then observed.

Figure 4.10 shows the corresponding percentage change in kinetic energy after the collision as a function of impact parameter. Significant gains in energy are seen at small impact parameters for the positive ion at low incident energy, consistent with the inaccurate integration of position. Energy is conserved to a better degree during the interaction with the negative ion. At higher incident electron energy, conservation of energy is better reproduced for both ion types. Thus, it may be expected that within an *ab initio* atomistic simulation, electron-electron and electron-acceptor interactions will be treated with sufficient accuracy while the energy gain during electron-donor interactions must be minimised.

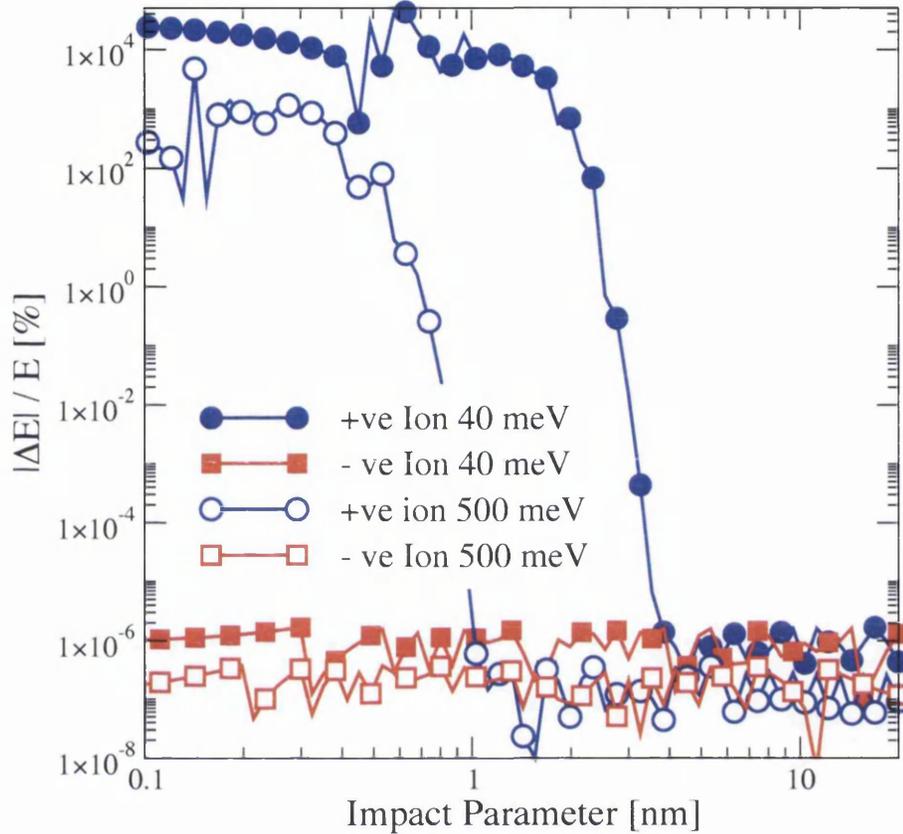


Figure 4.10: Magnitude of the percentage change in kinetic energy upon leaving the simulation domain. Results shown here include both positive and negative central charges and incident electron energies of 40 meV and 500 meV. Significant error in the final kinetic energy is seen for a positive charge at low incident electron energy while that of the negative charge remains in good agreement.

4.5 *ab initio* Ionized Impurity Scattering

The method described here for introducing *ab initio* ionized impurity scattering in mesh based particle simulations follows Hockney and Eastwood [118] who were among the first to introduce the technique [137, 138]. It is designed for the efficient simulation of a large system of charged particles where both long range and short range interactions must be resolved. It combines the direct force evaluation of the Particle-Particle (PP) technique, common to molecular dynamics and the simulation of small systems of parti-

cles, with the force interpolation of the Particle-Mesh (PM) technique, used to efficiently describe the long range interactions of a large ensemble of particles. The PP technique determines forces directly between pairs of particles through an analytical expression, thus accurately resolving the interaction. However, for large systems the direct evaluation of forces is computationally prohibitive. Problems also arise with the handling of materials with varying dielectric constants and devices with complex boundaries. Where it is sufficient to only consider long range interactions, significant improvement in efficiency, and the ability to simulate much larger ensembles, is achieved using the PM technique. In this case the field is obtained from the numerical solution of the Poisson equation using mesh based discretization. The field calculated at the mesh points is then interpolated to the current particle positions within the mesh. The drawback of the PM approach is the poor representation of the short range interactions. The Particle-Particle-Particle-Mesh (P³M) method combines the strengths of both techniques, using the mesh interpolated field to efficiently obtain long range interactions supplemented with a direct force evaluation over only the neighbouring particles to account for short range interactions.

Within this model, the net force on electron i may be given as

$$F_i = F^{mesh}(\mathbf{r}_i) + \sum_{j \neq i} F^{sr}(\mathbf{r}_i - \mathbf{r}_j) \quad (4.27)$$

where $F^{mesh}(\mathbf{r}_i)$ is the mesh force interpolated at the electron position \mathbf{r}_i , and $F^{sr}(\mathbf{r}_i - \mathbf{r}_j)$ is the short range correction for the force acting on electron i due to a neighbouring charge j , where the sum is over all such neighbouring charges. The mesh interpolated force was discussed previously in chapter 3, the remainder of this discussion is focused on the short range force correction.

4.5.1 Short Range Correction, F^{sr}

Using the CIC charge assignment scheme, the mesh based solution of the electrostatic potential associated with a point charge significantly underestimates the true Coulomb force within separations of approximately two mesh spacings [118]. For separations greater than this, the mesh interpolated forces are accurate to less than one percent compared with the analytical Coulomb force. Therefore, in order to achieve a good agreement with the exact Coulomb force, a force correction has to be applied for electrons separated from surrounding point charges by less than two mesh spacings. The short range radius r_{sr} defines a sphere centred about a particle and contains all the neighbouring charges for which the mesh force alone inadequately describes the force, and for which short range corrections must be applied. This radius will be dependent upon the form of the short range interaction and will be discussed later.

In general, the force contribution from the mesh associated with two charges separated by less than r_{sr} is not zero. This results in an overlap of the long-range mesh and short-range analytic force regimes. It is therefore necessary to define the short range correction as the difference between the analytic short range interaction $F^*(\mathbf{r}_i - \mathbf{r}_j)$, and the mesh resolved force in order to avoid double counting of forces. The short range correction is then expressed as

$$F^{sr}(\mathbf{r}_i - \mathbf{r}_j) = F^*(\mathbf{r}_i - \mathbf{r}_j) - F^{ref}(\mathbf{r}_i - \mathbf{r}_j) \quad (4.28)$$

where $F^{ref}(\mathbf{r}_i - \mathbf{r}_j)$ is the reference force and equals the mesh resolved force for two charges separated on the mesh.

4.5.2 Reference Force, F^{ref}

Although equation 4.28 defines the short range force through the difference of two terms which are solely functions of charge separation, in the case of the reference force this is

not accurate. The Poisson equation is solved on a Cartesian grid and the force extracted from the mesh based solution has the corresponding symmetry. For a given charge separation, the mesh force varies for different orientations of the charges with respect to the mesh. Similarly, dependent upon the charge assignment and force interpolation scheme employed, the mesh force differs with the position of the charges relative to the mesh due to the charges being spread over the mesh cells. The implied translational invariance and symmetry of the reference force in equation 4.28 is instead a 'best fit' approximation in one dimension. Accurately determining the reference force would require the mesh based solution of Poisson's equation to be known for all possible positions and orientations of a pair of charges within the volume defined by the short range radius and represents a significant increase in complexity of the problem. This issue has been somewhat addressed by tabulating the components of the mesh resolved force at steps of two angstroms throughout the volume associated with the short range force correction [139]. This introduces a very large memory overhead but allows the mesh force to be accurately described in three dimensions, yielding a more accurate force correction for different charge pair orientations and separations. However, the mesh force is based upon the solution of a point charge in a single position relative to the mesh. As such it still neglects the translational variation in the mesh force. Alternatively, an analytical interpolation scheme was proposed in order to describe the symmetry of the reference force within a cell and was reported with apparent success [140, 141]. However, this again makes assumptions about the position of one of the charges and also inadequately describes the translational variation within a cell.

The use of higher order charge assignment and force interpolation schemes can, through further smoothing and increasing of the short range radius, reduce the magnitude of the mesh force at short range. By minimising the short range mesh contribution at close separations, where the anisotropy is most significant, a more favourable description by

the reference force as solely a function of separation can be achieved. This in turn increases the accuracy of equation 4.28. In particular, TSC charge assignment and force interpolation, in conjunction with a four point finite difference scheme for evaluating the field at each mesh point, results in a force correction accurate to within one percent over both force regimes. Indeed, this is the method recommended and employed by Hockney and Eastwood where it was apparent that CIC charge assignment and force interpolation represented too great an error in both the magnitude and direction of the field within the short range region [118]. Unfortunately, higher order charge assignment and interpolation represents further computational overheads and, more importantly, larger finite sized particles which cause problems at abrupt heterojunctions, interfaces and simulation domain boundaries. For these reasons we accept the limitations and inaccuracies of using equation 4.28 with CIC charge assignment.

A suitable approximation of the reference force is then required for use in equation 4.28. An analytical approximation when using CIC charge assignment is given in [118] as

$$F^{mesh}(r) = \frac{1}{4\pi\epsilon} \begin{cases} \frac{1}{35r_{sr}^2} (224\xi - 224\xi^3 + 70\xi^4 + 48\xi^5 - 21\xi^6) & 0 \leq \xi \leq 1 \\ \frac{1}{35r_{sr}^2} (12\xi^{-2} - 224 + 896\xi - 840\xi^2 + 224\xi^3 + 70\xi^4 - 48\xi^5 + 7\xi^6) & \frac{r_{sr}}{2} < \xi \leq 2 \\ \frac{1}{r^2} & \xi > 2 \end{cases} \quad (4.29)$$

where

$$\xi = \frac{2r}{r_{rs}} \quad (4.30)$$

This analytical form has reportedly been used with success in describing low field bulk concentration dependent mobilities in silicon, and it is also argued that it has an advantage in being simply extendable for application within a non-uniform mesh [79].

For all simulations presented here, the reference force is obtained via the alternative

method described by Gross [77]. In this approach the reference force is evaluated directly from the mesh based solution of a single charge, where the mesh used has the same dimensions as that used in the subsequent device simulation. The point charge is placed in the middle of a cell where the solution has the highest degree of symmetry. This symmetry attempts to minimise errors in treating the mesh force as spherically symmetric. The mesh force is then interpolated through the centre of the cell, parallel to a principal axis, and tabulated in steps of half an angstrom. This is similar to the previously described 3D approach [139] in that the actual mesh force is used to tabulate the force correction term. However this method only tabulates the correction in 1D and as such requires far less memory overhead for storing the tabulated values. This method is therefore very simple and efficient to implement but can not be applied to a non-uniform mesh. Since all simulations presented here are performed on a uniform mesh, this is not a concern.

4.5.3 Analytic Short Range Force, F^*

Figure 4.11 shows the mesh interpolated field for the solution of a single point charge placed in the middle of the mesh cell in comparison with the corresponding true Coulomb field. The mesh spacing is 5 nm . Agreement is good at long range, with the error in the mesh force increasing at close separation. Also plotted is the mesh force with the short range correction. The comparison is made for three different orientations defined by the charge and distinct points of symmetry of the cubic mesh cell. This is to highlight the variation in respect of the adopted spherically symmetric mesh force. The first direction, labelled 1D in figure 4.11, is parallel to the x -axis and is identical to the approach used in determining the reference force, thus exactly reproduces the short range field. The second direction, labelled 2D, is a line in the xy plane making an angle of 45 degrees with the first, while the direction labelled 3D passes through the corner node towards the charge. The

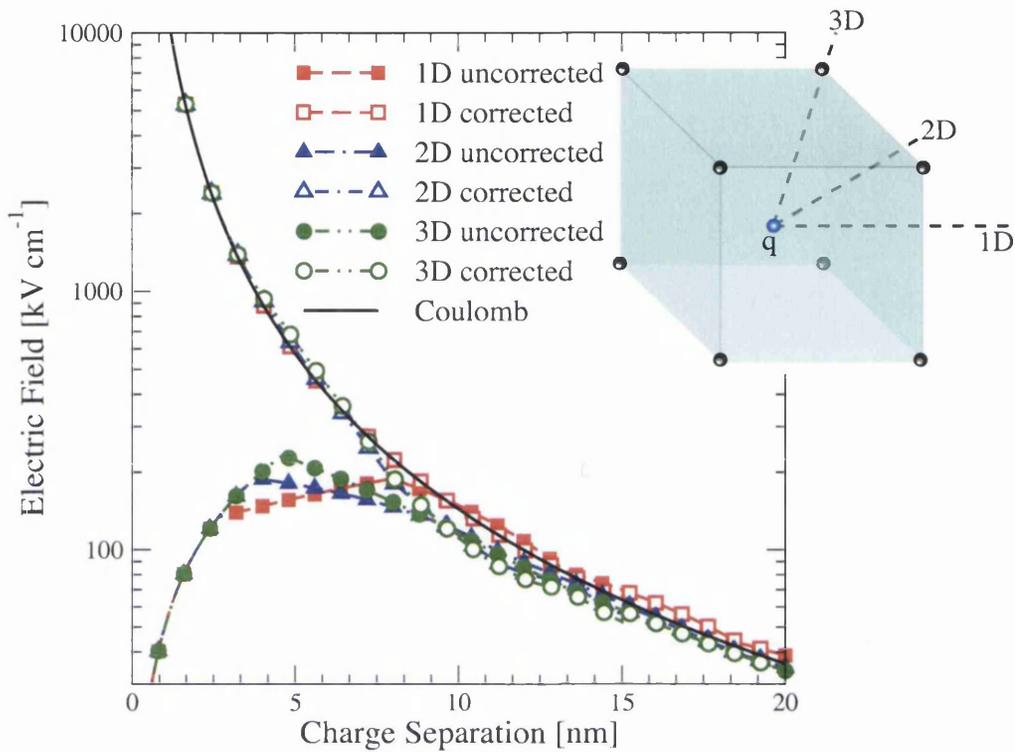


Figure 4.11: Comparison of the mesh interpolated and analytical fields. Agreement is seen at long range while the mesh underestimates the analytical field at close separations. The mesh corrected field, also shown, is seen to agree with the analytical field at close separations while the transition between the regimes is marked by position dependent variation.

later reproduction of bulk mobilities via *ab initio* simulation justifies the error introduced by this simple approach in the transitional region between the long range mesh regime and the analytical short range regime.

In the earlier reproduction of Coulomb scattering, significant unphysical gains in energy were seen in propagating a charge in the rapidly varying field associated with an attractive Coulomb potential. This error needs to be minimised and such minimisation is achieved by defining an alternative short range interaction. Three models were initially tested by Gross [77] and are plotted in figure 4.12. Each model follows the Coulomb

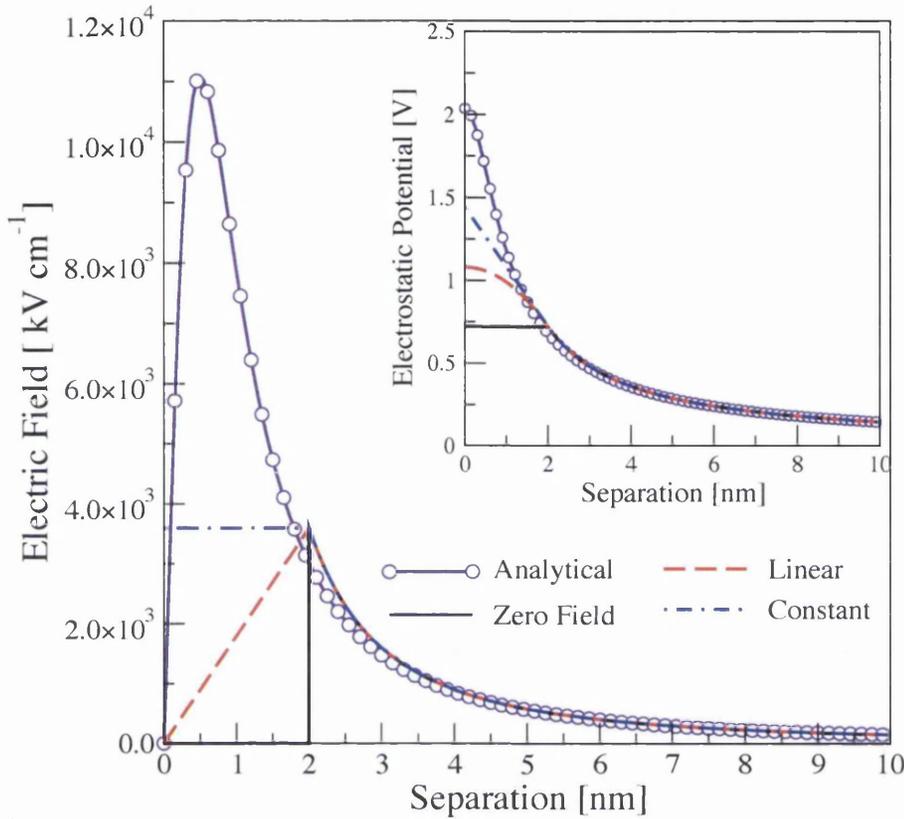


Figure 4.12: Potential and electric field profiles for three proposed short range interaction models to minimise errors in propagation [77]. Also shown is a fourth analytical model adopted here.

force up to a separation of 2 nm , which is approximately the ground state radius of the hydrogenic model donor in silicon. Below this radius one of the following three approximations is used; (i) a constant field equal to the field at 2 nm , (ii) a constant field equal to zero, or (iii) a field linearly decreasing to zero from the value at 2 nm . Of the three approximations, the linearly decreasing field was reported to give the best results in device simulation. An alternative analytical form is used here, proposed in [118]

$$F(\mathbf{r})^* = \frac{e\mathbf{r}}{4\pi\epsilon_0\epsilon_r \left(\mathbf{r}^2 + \frac{1}{2}r_c^2\right)^{\frac{3}{2}}} \quad (4.31)$$

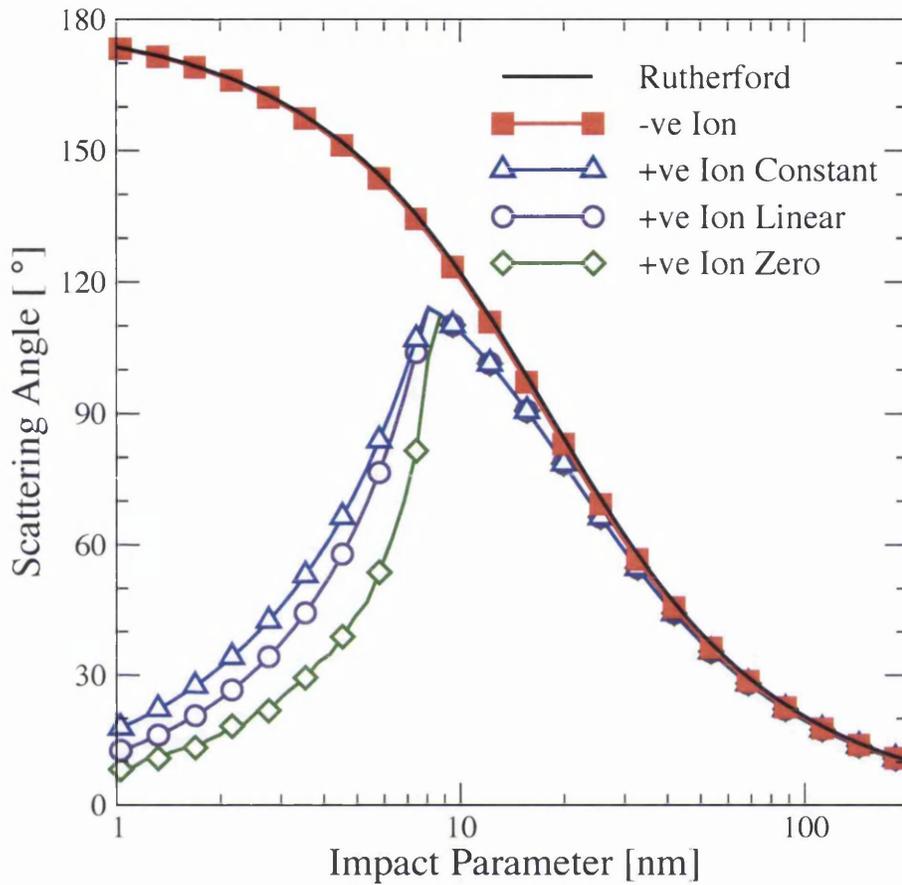


Figure 4.13: Scattering angle dependence upon impact parameter using three short range field models [77]. Scattering follows the Coulomb case but rapidly drops at smaller impact parameters.

with the corresponding potential given by

$$V(\mathbf{r})^* = \frac{e}{4\pi\epsilon_0\epsilon_r \left(\mathbf{r}^2 + \frac{1}{2}r_c^2\right)^{\frac{1}{2}}} \quad (4.32)$$

This matches the Coulomb law at large separations, reaches a maximum value at the cut-off radius r_c and approaches zero as the separation \mathbf{r} approaches zero. The approximation is qualitatively similar to the model adopted in [77] and is plotted for comparison in figure 4.12 where r_c is taken equal to 0.5 nm . The choice of an analytical model covering the

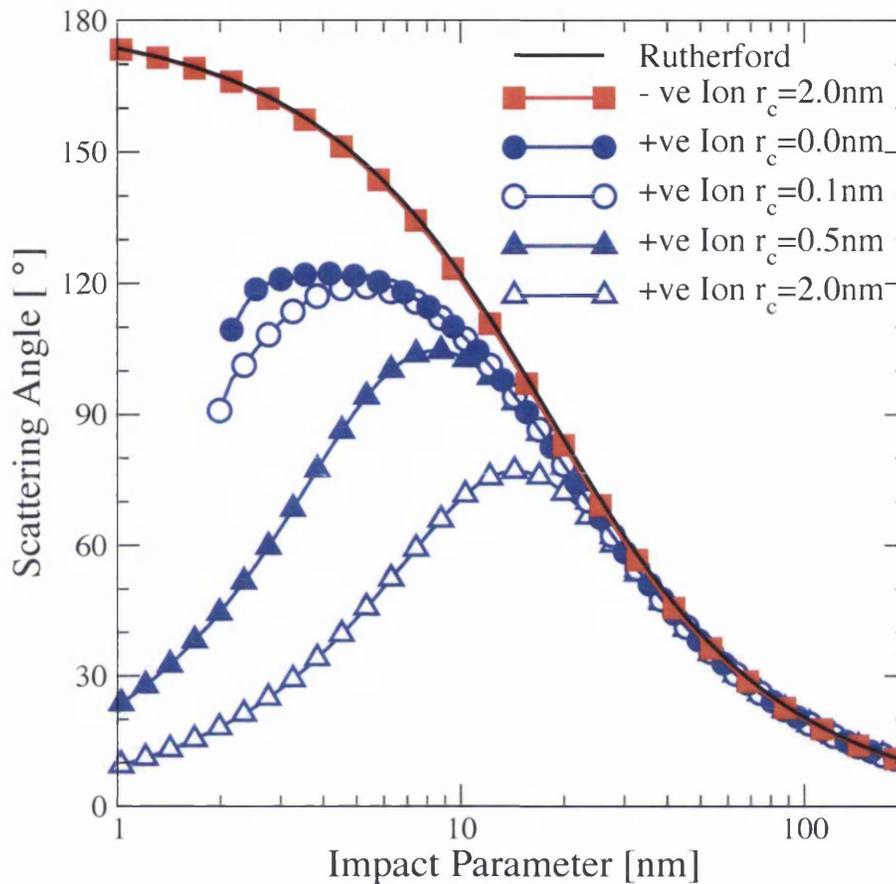


Figure 4.14: Comparison of scattering angle versus impact parameter for the modified potential with variety of cut-off radii at an incident electron energy of 40meV . The interaction with a negative ion again follows the Rutherford result while larger cut-off radii favour move forward scattering.

entire range of the short range interaction is preferential to the three abrupt short range models since both the field and the derivative of the field are continuous and lead to better conservation of energy during propagation.

Limiting the short range interaction in order to better conserve energy will limit the ability of a point charge to act as a true Coulomb scattering centre. To assess the trade-off between energy conservation and scattering angle dependence, the simulation of Coulomb scattering of electrons in section 4.4.3 is repeated using the four short range force models

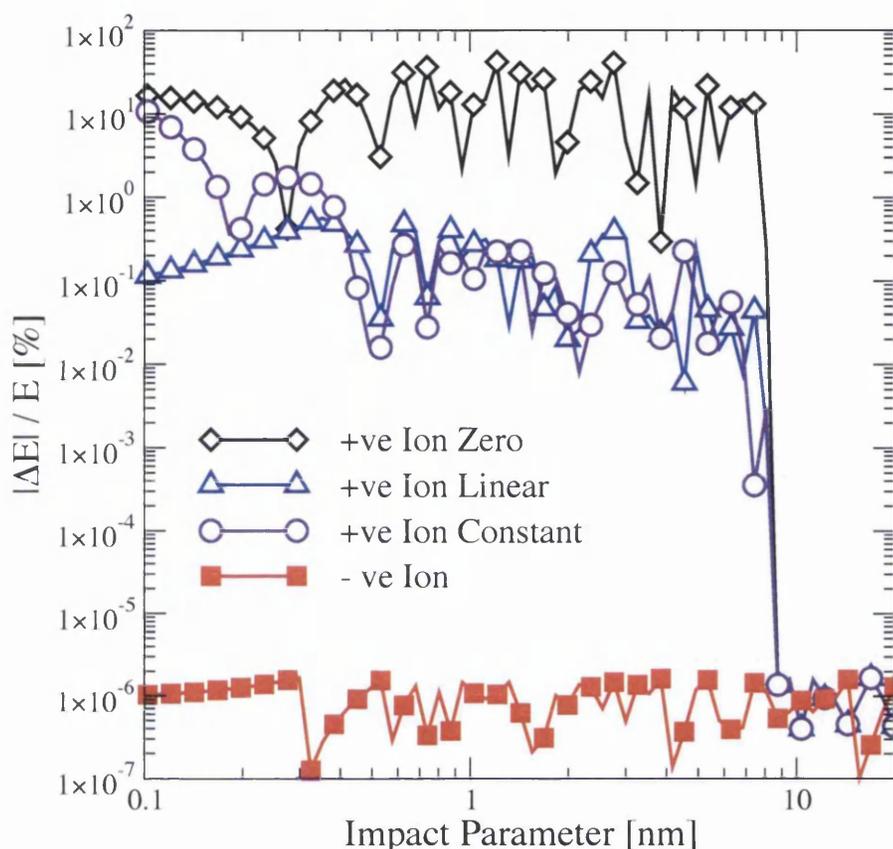


Figure 4.15: Percentage change in kinetic energy after the collision for the three short-range field models tested by Gross. Of the three models, the linearly decreasing field shows marginally better energy conservation with a maximum error in the kinetic energy of less than 1 percent.

discussed above. Figure 4.13 shows the scattering angle dependence at an incident energy of 40 meV resulting from the three models in [77]. In all cases, scattering from the negative ion remains accurate since modification of the Coulomb law is made sufficiently near the ion that electrons don't see it at this energy. For the positive ion, the scattering angle is maintained for impact parameters down to 9 nm where it rapidly falls due to the reduced interaction close to the ion. This results in a maximum scattering angle of nearly 120 degrees for any interaction. Figure 4.14 shows similar results using the modified force of equation 4.31 for values of the cut-off radius r_c equal to 2.0 nm , 0.5 nm , 0.1 nm and

0.0 nm . Again negative ions are accurately treated, even for the largest cut-off, while the maximum scattering angle due to a positive ion is reduced for larger cut-off radii.

The percentage change in energy is shown for the three models in [77] in figure 4.15. Energy is poorly conserved, though there is some improvement with respect to the pure Coulomb potential (see figure 4.10). The linear short range model shows the best conservation of energy, in accordance with the claim made in [77], with a typical gain in energy greater than 0.1% . The resulting percentage change in energy using the analytical model with cut-off radii equal to 2.0 nm , 0.5 nm , 0.1 nm and 0.0 nm are shown in figure 4.16. The model provides energy conservation orders of magnitude better than the alternative short range models for cut-off radii equal to 0.5 nm and below.

Using a cut-off radius of 0.5 nm within the short range model described by equation 4.31 results in a force and potential profile similar to the model adopted in [77] up to 2 nm . The field then exceeds the field corresponding to the other approaches, reaching a maximum before falling to zero. As the smoother form of the analytical short range model shows significant improvement in energy conservation and a similar scattering angle dependence compared with other models, the model was adopted in the rest of this work. The imposed limiting of the point charge interaction is not completely without physical basis since an ionised impurity is not a point charge. Interaction with the remaining valence and inner shell electrons will reduce the pure Coulomb potential at short distances from the nucleus. However the empirical model adopted pragmatically captures the main effects of scattering without attempting to accurately reproduce the real impurity potentials.

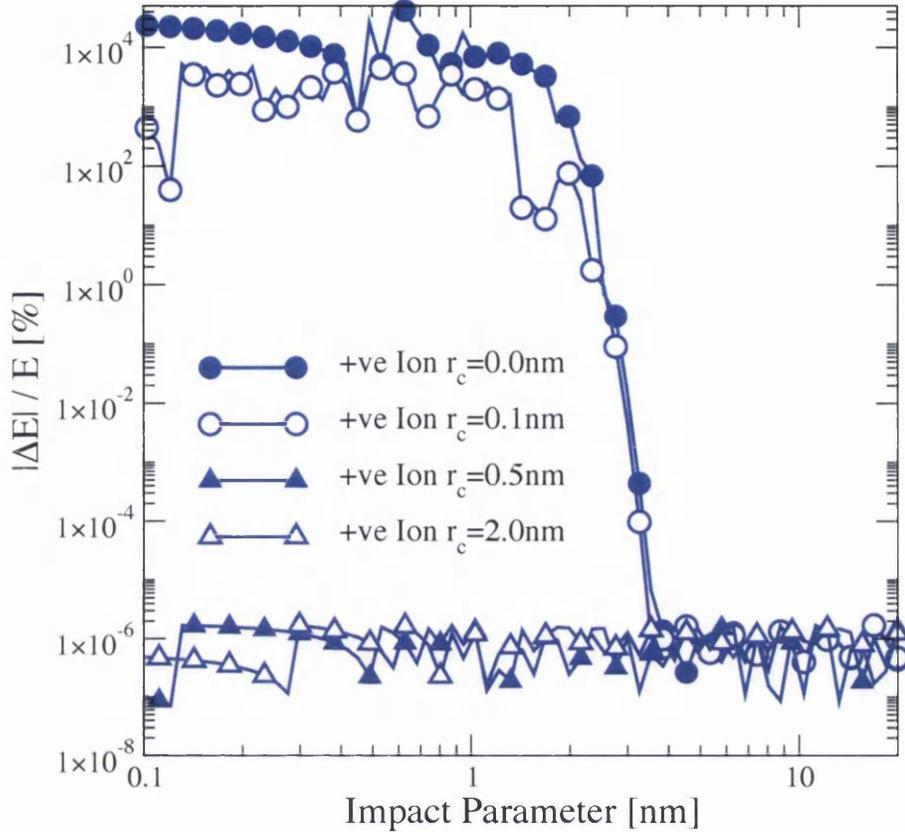


Figure 4.16: Magnitude of the percentage change in kinetic energy over the collision for a positive central charge using the modified potential with a range of cut-off radii. Improved conservation of energy is seen for larger cut-off radii.

4.5.4 Short Range Radius, r_{sr}

The mesh field accurately resolves the analytical Coulomb field at separations beyond two mesh spacings. However, the use of the analytic short range force model in the previous section results in a deviation from the Coulomb field at short range. For the simulation of ultra-small devices with a finely resolved mesh in this work, it is not guaranteed that the adopted short range force will agree with the mesh force at distances larger than two mesh spacings. The short range radius r_{sr} over which corrections are performed has been re-defined to be the minimum distance at which the form of the short range interaction

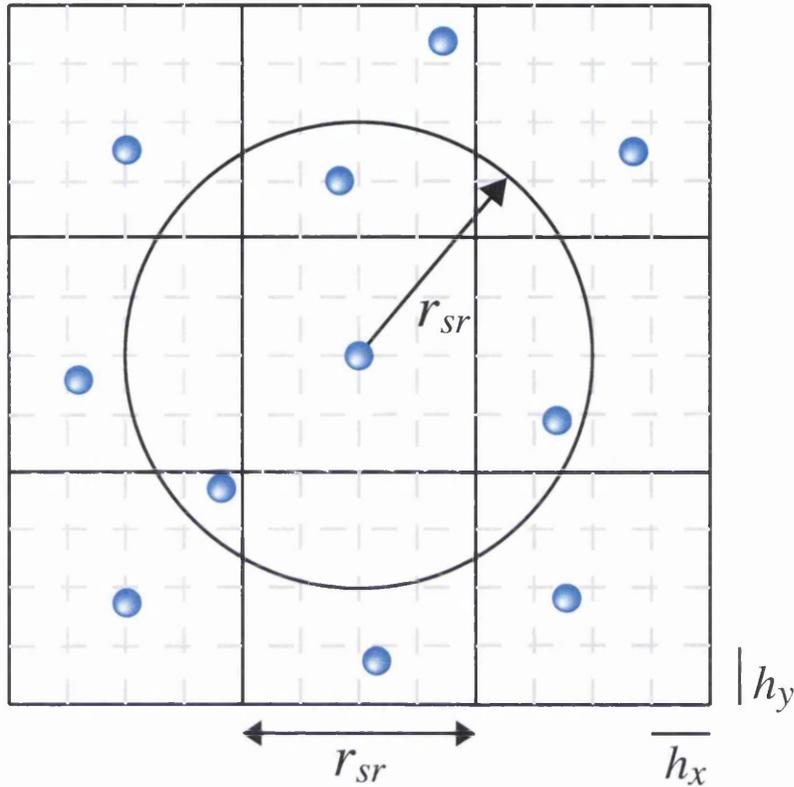


Figure 4.17: The force correction is evaluated for those electrons separated by $\leq r_{sr}$. Separations are determined only for those electrons stored in adjacent chaining mesh cells, limiting the required number of comparisons.

agrees with the point charge potential on the simulation mesh. A value of 15 nm is used for r_{sr} in simulations using mesh spacings smaller than 7.5 nm , otherwise r_{sr} is equal to twice the cell width. This guarantees that the correction matches on any mesh.

4.5.5 Application of Force Correction

In order to avoid calculating the separation between every pair of particles within the simulation domain to determine whether or not they fall within r_{sr} , a time consuming process deliberately avoided by utilising the mesh force, particles are referenced to a second, coarse, mesh termed the chaining mesh. The chaining mesh is a uniform mesh with cell

width equal to r_{sr} . A linked list sorts the particles within each chaining cell following the discussion given in Hockney and Eastwood [118]. The calculation of the distance between neighbouring particles is now limited to those in the surrounding chaining cells as illustrated in figure 4.17. The number of comparisons required then increases with the short range radius and the particle concentration at any point, but is significantly less than in the brute force alternative.

Incorporating the short range force correction recovers the full charge interaction that the mesh alone can not resolve. Care must be taken to ensure that the correction of the mesh resolved charge interaction within a simulation is physically accurate. Incorporation of the full electron-electron interaction within a continuously doped simulation leads to the dominance of the inter-electron forces and results in an extension of the depletion regions and an increase in the potential drop across them. Within an atomistic simulation, both electron-electron and electron-ion interactions must be resolved consistently to account for charge screening. Inclusion of the full electron-donor interaction without the electron-electron interaction affects the electron screening of the donors and a results in an un-physically high electron concentration localised at donor sites. Therefore, within self consistent atomistic simulations both the electron-electron and electron-ion interactions should be included together. Within continuous doping simulations, the electron-electron short range force correction should not be used.

4.6 *ab initio* Simulation of Doping Concentration Dependent Mobility in Silicon

Having determined both a form for the short range charge interaction that maximises scattering while minimising errors in energy conservation and a method for applying the short

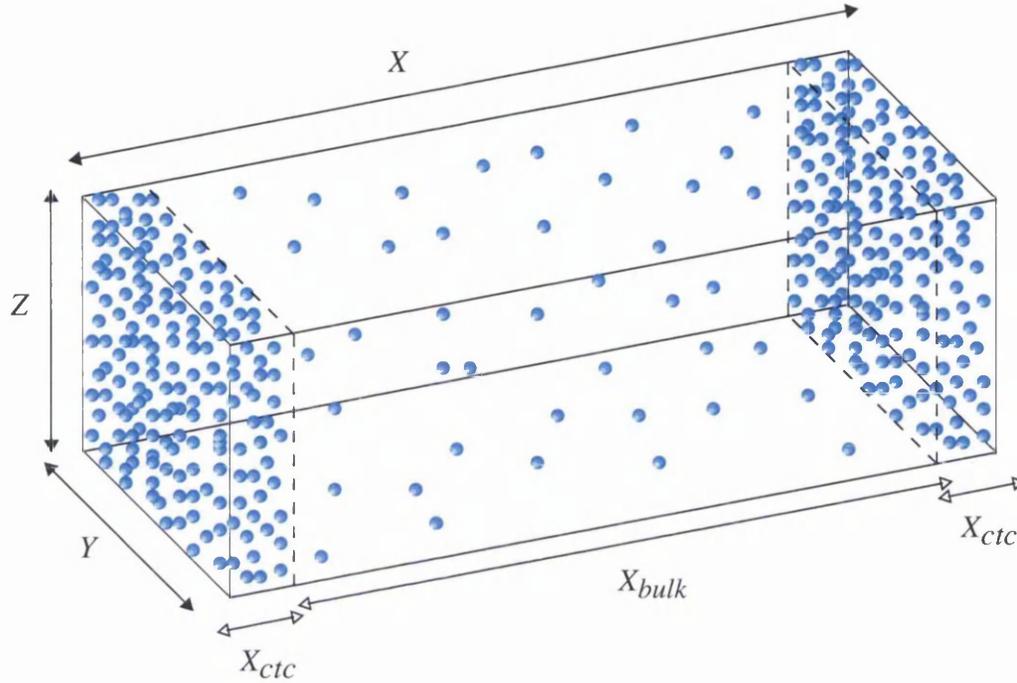


Figure 4.18: Schematic 3D atomistic simulation structure for recovery of bulk mobility. The central doped region, length X_{bulk} , is the concentration under investigation and is situated between two more heavily doped contact regions, length X_{ctc} .

range force correction, it is important now to validate the model, demonstrating that it can accurately describe ionized impurity scattering. Therefore in this section, simulations are performed to recover the well known bulk concentration dependent electron mobility in silicon over a range of doping concentrations.

4.6.1 Bulk Simulation Structure

The evaluation of the doping concentration dependence of the bulk mobility has been attained by the simulation of a simple resistor structure. Each simulated resistor has a high doping concentration in the contact regions, relative to the central region in which the mobility is evaluated, in order to closely maintain ideal Ohmic contact behaviour.

X (nm)	Y, Z (nm)	X_{ctc} (nm)	X_{bulk} (nm)	N_{ctc} (cm^{-3})	N_{bulk} (cm^{-3})	ΔV (V)
14080	3520	880	12320	1×10^{15}	1×10^{14}	1.2320
6400	1600	400	5600	1×10^{16}	1×10^{15}	0.5600
2944	736	184	2576	1×10^{17}	1×10^{16}	0.2576
1408	352	88	1232	1×10^{18}	1×10^{17}	0.1232
640	160	40	560	5×10^{18}	1×10^{18}	0.0560

Table 4.1: Dimensions and doping concentrations of the different regions within the resistor devices and the applied potential difference.

A potential difference is applied across the device which drops across the central region resulting in a low electric field of around 1 kV cm^{-1} . Devices with varying doping concentrations in the central part of the resistor are all simulated on a uniform mesh consisting of $129 \times 33 \times 33$ mesh points, with physical dimensions such that there are approximately 10,000 donors within each contact region and 15,000 donors within the central region. The general device structure is illustrated in figure 4.18 while specific dimensions and doping concentrations are presented in table 4.1.

Self consistent EMC simulations were performed in which the short range force corrections for both the electron-electron and electron-donor interactions were included. The time step for integration and successive solutions of Poisson's equation was set at 0.1 fs to accurately resolve the short range charge interactions. Simulations were performed over a simulated time of 6 ps , allowing 2 ps for the decay of initial transients, at a temperature of 300K .

For each device, the average electric field within the central region was obtained from the time averaged mesh potential throughout the region and was used in combination with the average electron velocity in the region to estimate the mobility. The large numbers of donors within the region approximate the self averaging donor distribution in bulk material and so the obtained mobility value is reported as the bulk mobility value.

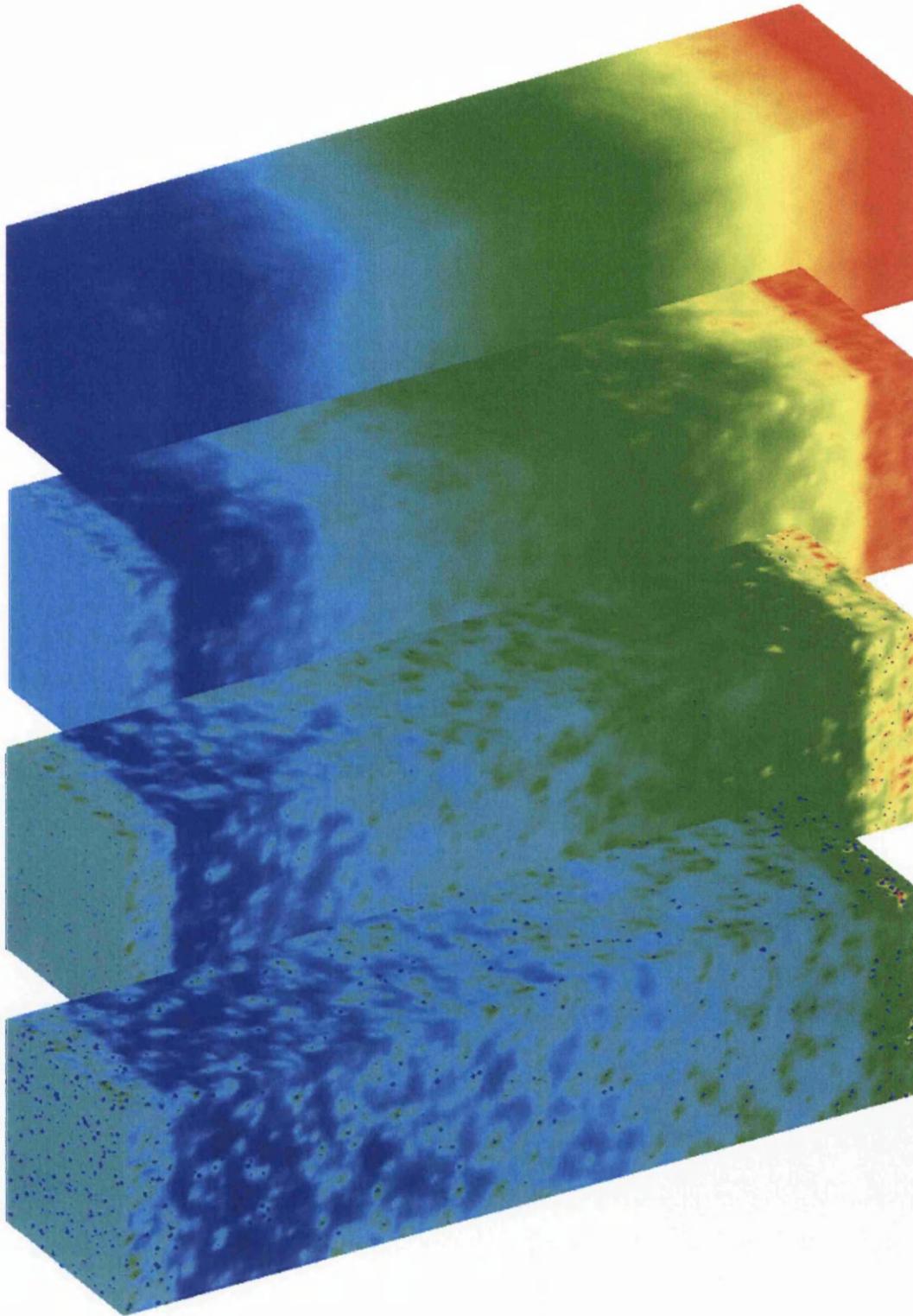


Figure 4.19: Potential throughout simulated structure for the extraction of bulk mobility for donor concentrations 10^{15} cm^{-3} , 10^{16} cm^{-3} , 10^{17} cm^{-3} , and 10^{18} cm^{-3} from top to bottom respectively

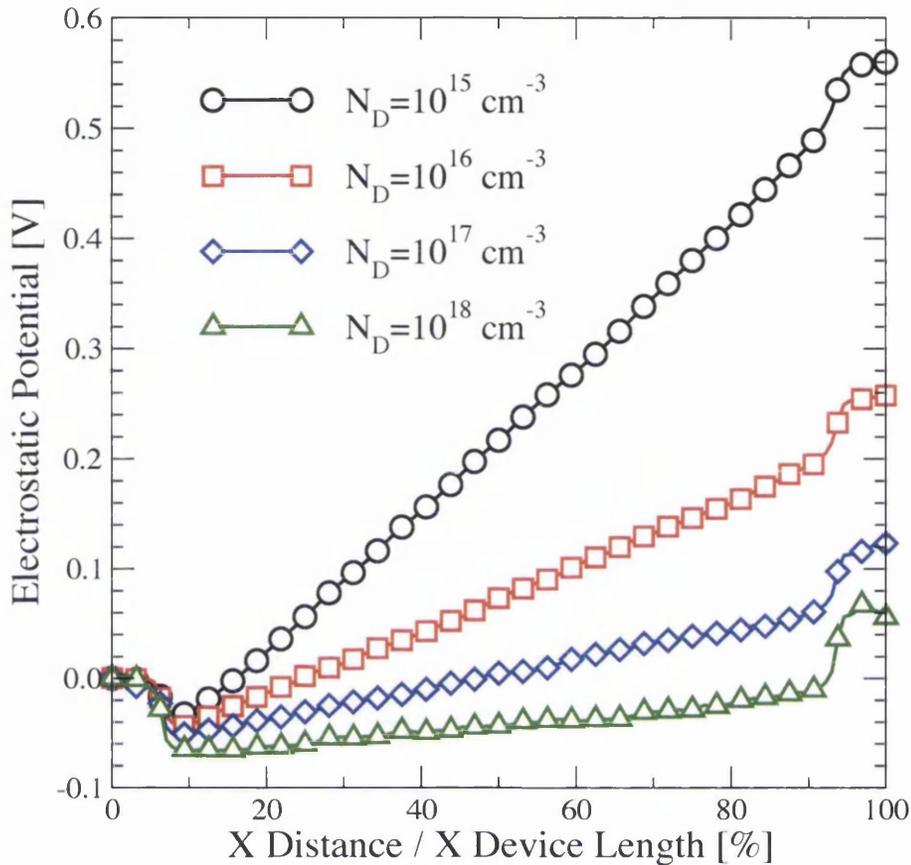


Figure 4.20: Potential profile normalised along the x -axis for each device. A linear profile in the central region is seen for each device with more variation apparent in smaller, higher doped devices.

4.6.2 Bulk mobility results

Electrostatic Potential and Electron Concentration Figure 4.19 shows the 3D steady-state potential throughout the resistor structures corresponding to the simulation of bulk mobilities at donor concentrations of $N_D = 1 \times 10^{15} \text{ cm}^{-3}$, $1 \times 10^{16} \text{ cm}^{-3}$, $1 \times 10^{17} \text{ cm}^{-3}$ and $1 \times 10^{18} \text{ cm}^{-3}$. The position of the individual donors are marked and are clearly seen to correspond to the local variations in the potential which are better resolved for the simulations at higher doping concentrations due to the use of the finer mesh.

The steady state averaged potential distributions along the resistor length are shown

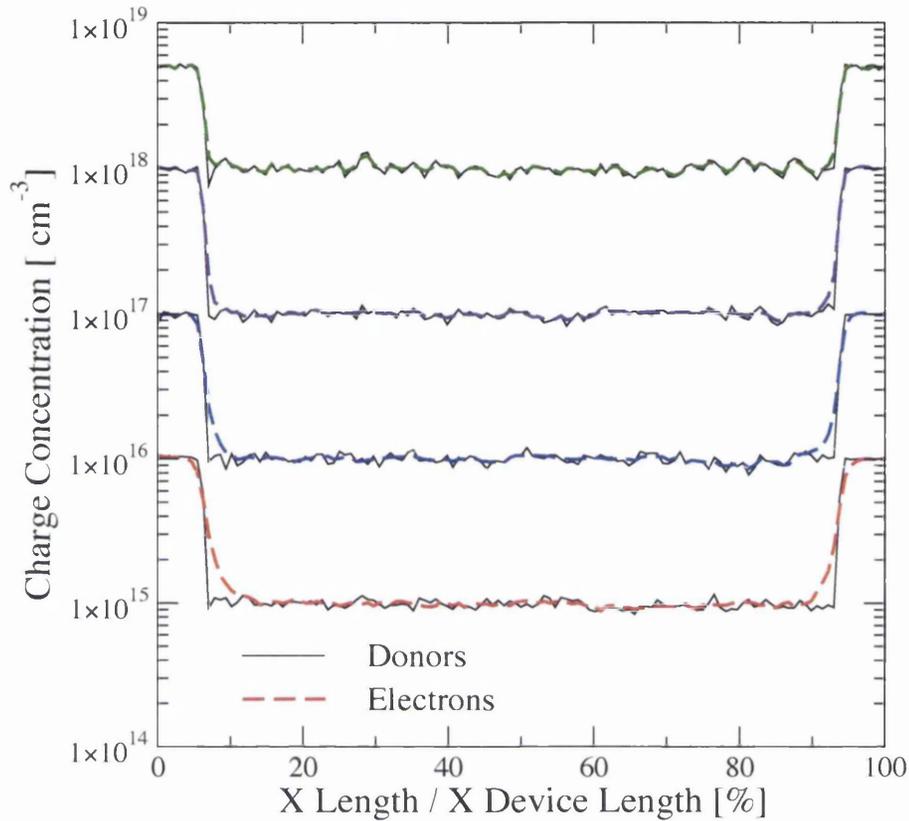


Figure 4.21: Atomistic donor profile and electron concentration along the normalised x -axis for each device.

for the same devices in figure 4.20, where the device lengths have been normalised to fit the same graph. Similarly, the averaged electron concentration distributions along the resistor are shown in comparison with the averaged donor concentration distribution in figure 4.21 for the same set of devices, again with the length of each device normalised. Good maintenance of the contacts is noted in all cases and the electron concentration is seen to closely follow the donor concentration. This is in turn reflected in the flat potential in the contact regions and the linear drop, with variations due to atomicity, in the central region.

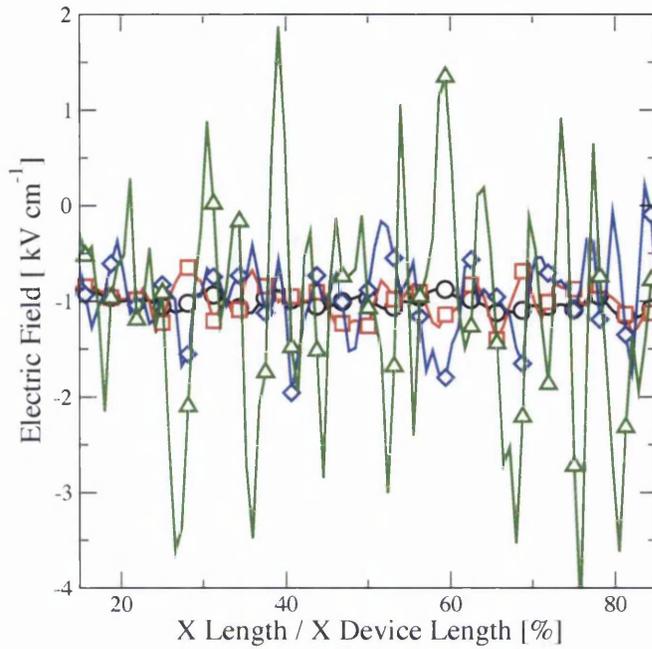


Figure 4.22: Electric field profile within the listed simulated devices. In all cases the average field is close to 1 kV cm^{-1} . Legend is same as in figure 4.20

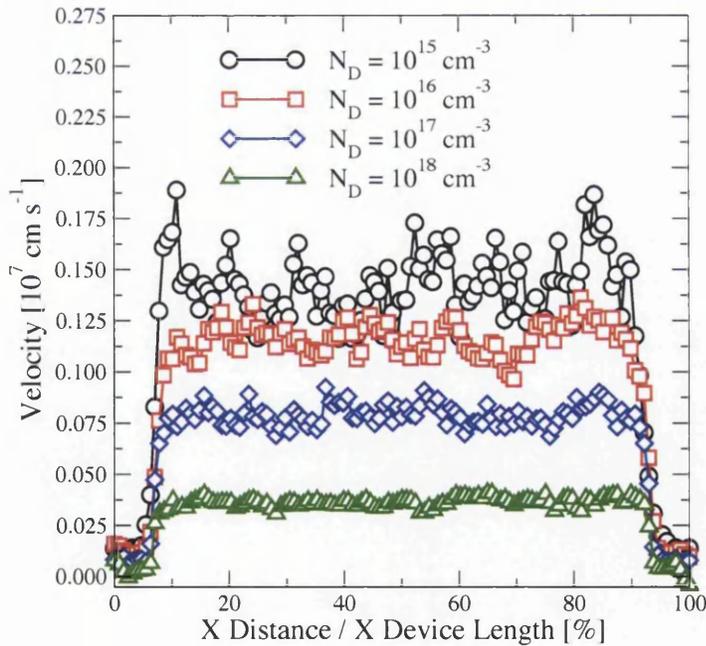


Figure 4.23: Average electron velocity profile within each device, normalised along the x -axis

v ($\times 10^7$ $cm s^{-1}$)	E ($kV cm^{-1}$)	μ ($cm^2 V^{-1} s^{-1}$)
0.1424	1.007	1414
0.1160	1.002	1157
0.0790	0.982	804
0.0339	1.104	307

Table 4.3: Average electric field and electron velocity throughout the central device region. Mobility as calculated from these values is tabulated and is also plotted in figure 4.24

Electric Field and Electron Velocity The average electric field, obtained from the average electrostatic potential throughout the device, is shown for each of the simulated devices in figure 4.22. The average value of the field in the central region, listed in table 4.3, is close to $1 kV cm^{-1}$ in all cases. Variation in the field is seen at higher doping concentrations due to the discrete charges being more accurately resolved in the mesh force.

The corresponding average electron velocity throughout the devices is shown in figure 4.23 and the average value throughout the central region is also given in table 4.3. Given the average electron velocity and the average field in the central region, the bulk mobility can be approximated by the ratio of the two.

Bulk Mobility Figure 4.24 shows the simulated bulk mobility from *ab initio* simulation compared to experimental values. Also shown is the mobility values obtained from impurity scattering rates as reported in chapter 3 and simulation results from *ab initio* simulation presented in [55, 77]. Excellent agreement is seen over the simulated range in doping concentration, better than compared with the results reported in [77]. It is also interesting to note that the *ab initio* approach better describes ionized impurity scattering

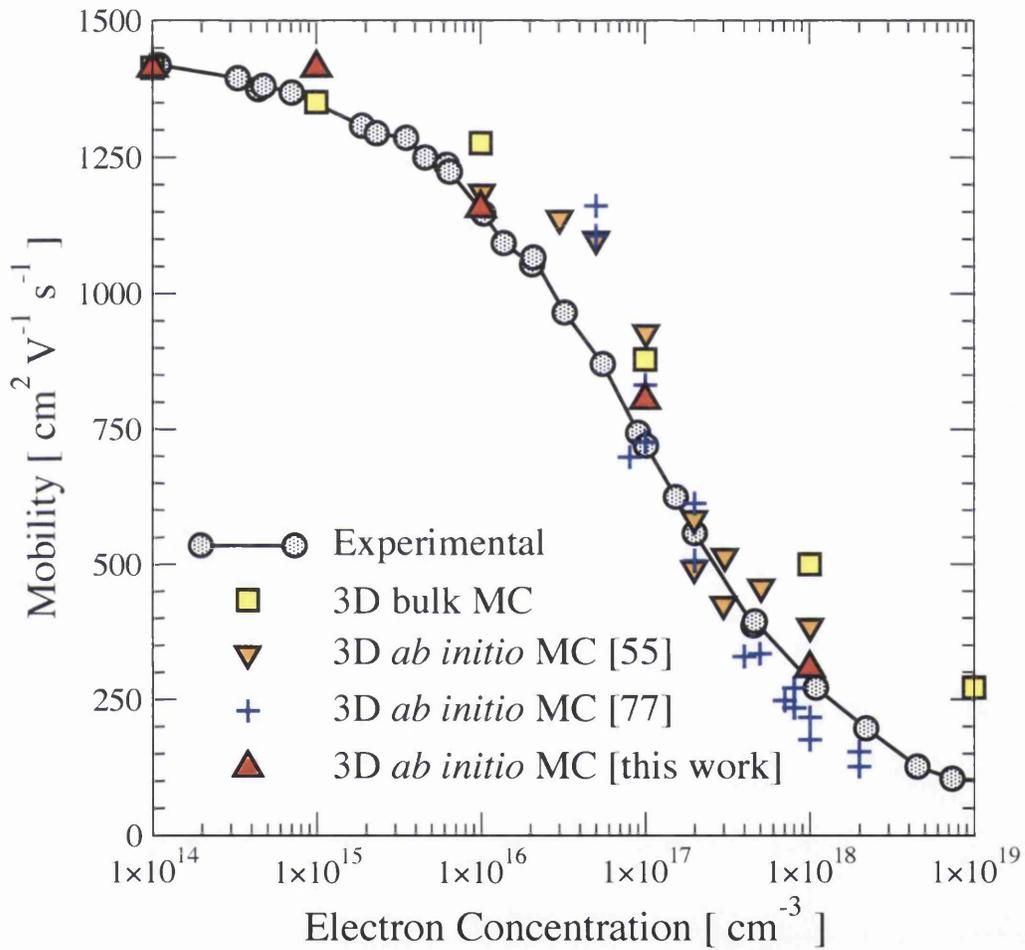


Figure 4.24: Reproduction of low field concentration dependent bulk electron mobility in silicon at 300K from 3D *ab initio* simulation. Comparison with experimental data, BH scattering rate and simulated mobility data from [55] and [77] is also shown.

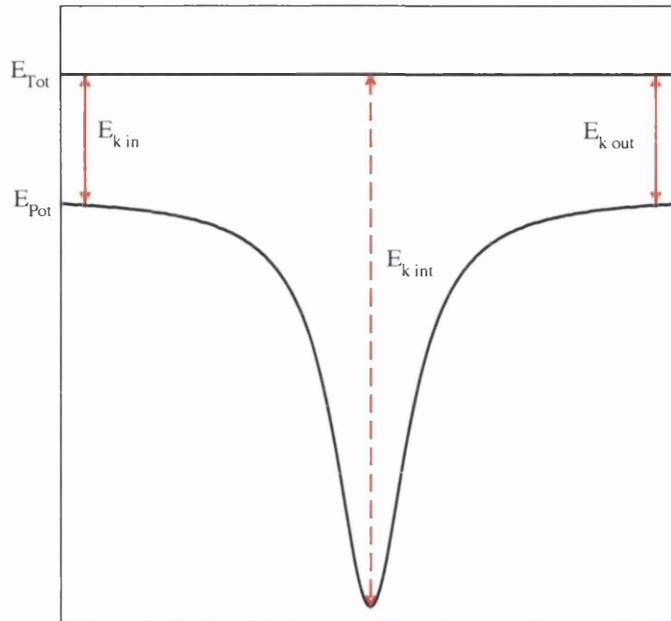


Figure 4.25: Representation of total energy, kinetic energy and potential energy plotted against position for an interaction with an ionised impurity. Energy is conserved throughout, but kinetic energy increases in the vicinity of the attractive potential.

compared with the Brooks-Herring approach.

Ensemble Average Energy Published results of *ab initio* simulated bulk mobility have to date only shown the calculated mobility values with no supporting evidence demonstrating accuracy in the propagation. Since unphysical heating can arise from poor integration of the equations of motion in the impurity potential, this can be misleading as the ensemble average velocity may agree with the bulk velocity but the distribution in velocities over the ensemble may be un-physically large. However, a measure of the conservation of energy during the simulation is hard to obtain. Figure 4.25 shows the conservation of total energy during the interaction with an attractive potential. An electron initially has kinetic energy E_{kin} and after the interaction has identical energy E_{kout} , but the kinetic energy during the interaction E_{kint} increases as the electron loses potential

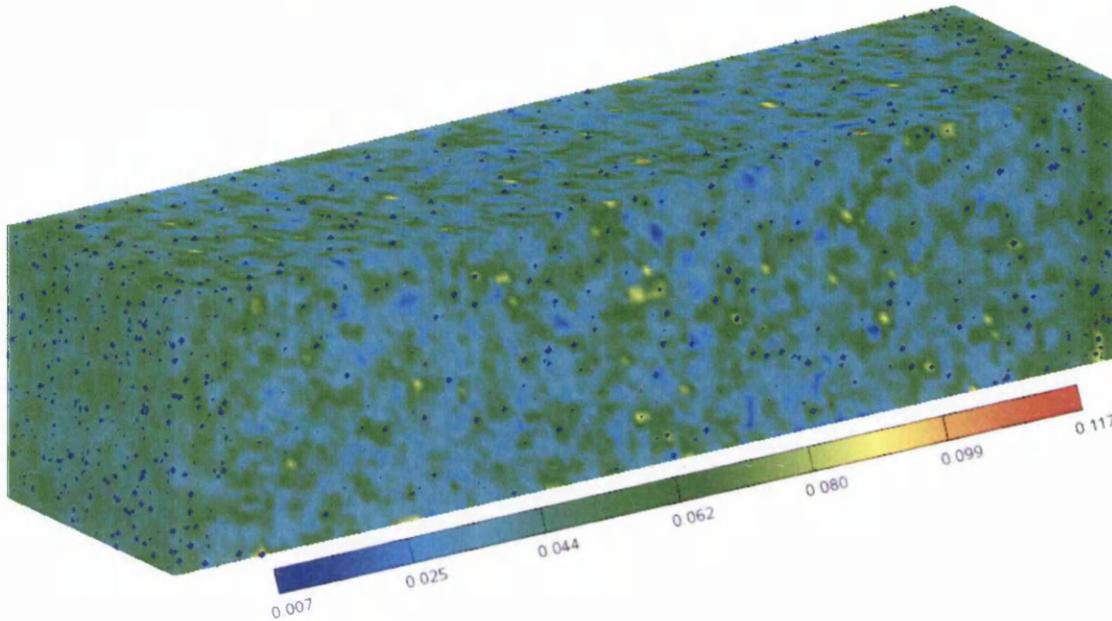


Figure 4.26: Electron kinetic energy distribution throughout structure simulating bulk mobility for $N_D = 1 \times 10^{18} \text{ cm}^{-3}$. Local increases in kinetic energy are observed around the marked donor positions. Energy measured in eV

energy before returning to its original value. The total energy is always conserved. The energy in the MC simulation refers to the electron kinetic energy and as such it varies during impurity scattering. The calculated steady-state ensemble average electron energy is the average kinetic energy over the ensemble averaged over time, and thus is expected to increase with donor concentration due to more frequent scattering and increases in kinetic energy. This increase in average energy with doping concentration is not direct evidence of artificial heating through inaccurate propagation.

The electron temperature corresponding to the simulation of the lowest donor concentration of $N_D = 1 \times 10^{15} \text{ cm}^{-3}$, is 303K. This is just 1% above the lattice temperature of 300K and is attributed to the energy gained by the applied field. The highest distribution temperature corresponds to the simulation at the highest donor concentration and is 342K. That this is not due wholly to some unphysical energy gain can be seen by looking at the

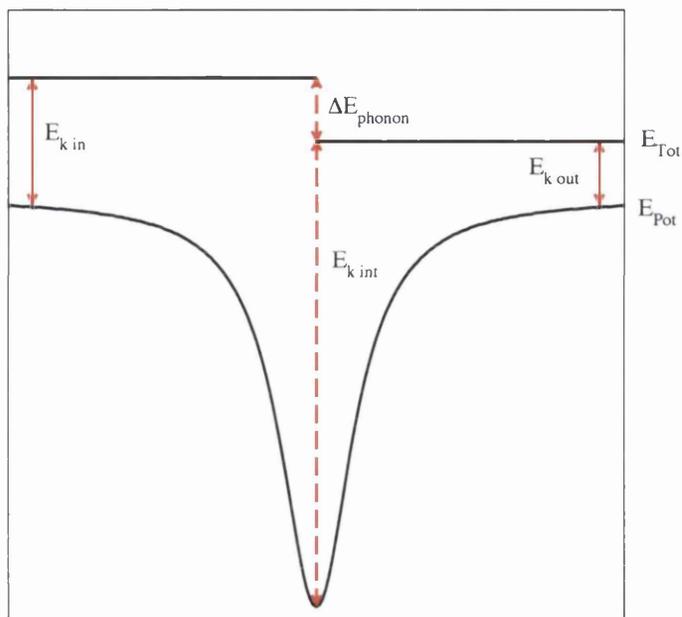


Figure 4.27: Representation of total energy, kinetic energy and potential energy plotted against position for an interaction with an ionised impurity. Increased kinetic energy makes more probable the emission of a phonon of energy E_{phonon} and loss of energy.

average electron kinetic energy distribution in real space. This is shown for the device used to simulate the doping concentration $N_D = 1 \times 10^{18} \text{ cm}^{-3}$ in figure 4.26. The local increase in kinetic energy surrounding the location of individual impurity atoms can clearly be seen. This is an indication that energy is accurately conserved within the simulation, while the increased temperature of the energy distribution is associated with the time sampling of electrons encountering more frequent short range interactions.

Measurements of the total energy of the electrons is complicated by phonon emission and absorption which does not conserve the electron's total energy. In this regard it is interesting to note that the probability for phonon emission is higher with increasing electron kinetic energy and thus the probability for phonon emission is relatively high during the ionised impurity scattering process. This is illustrated in figure 4.27. Electrons may emit phonons, reducing their total energy by amount ΔE_{phonon} , and this can provide

a means of trapping electrons around impurities if sufficient energy is lost. This should however remain a rare event since the impurity interaction time is relatively short.

4.7 Atomistic N-I-N Diode

In Chapter 3 the functionality of the developed 3D Self Consistent EMC simulator using traditional continuous doping was demonstrated through the simulation of an N-I-N diode. In order to further validate the application of the *ab initio* approach to impurity scattering within an atomistic simulation, the same N-I-N diode is simulated here with the modification that the continuous doping is replaced by a randomly generated distribution of singly ionized donor atoms.

The device contains 1,139 donors randomly placed within the simulation domain according to the doping concentration. This represents significantly fewer donors than previously simulated and variation from the bulk properties may be expected, but similarities between the continuous and atomistic doping results should remain due to the large size of the device and the dominance of phonon scattering.

4.7.1 'atomistic' electrostatic potential and electron concentration

Figure 4.28 shows the 3D distribution of the electrostatic potential obtained from the simulation of the atomistic N-I-N diode (cf. figure 4.28). The individual donors are marked and can again be seen to coincide with local increases in potential while the slowly varying potential profile is similar to the continuous device. The agreement between the slowly varying average potential within the atomistic device and the continuous potential is illustrated in figure 4.29. A slight difference in the depletion regions in the atomistic case is associated with variation in the non uniform doping distribution compared to the continuous case.

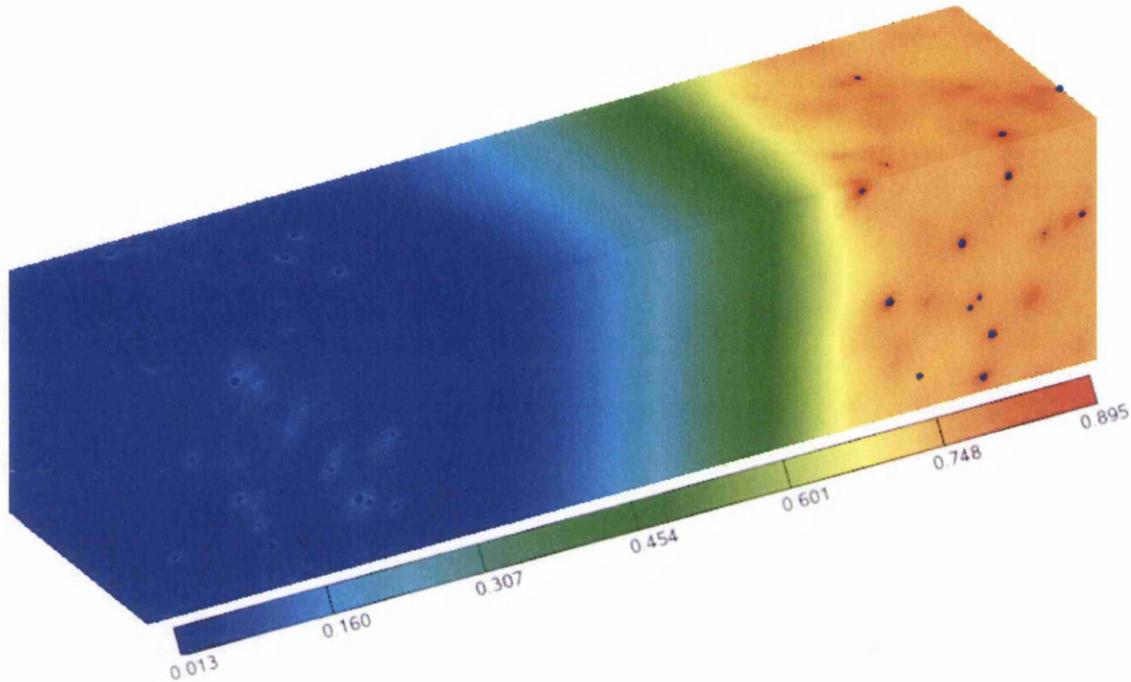


Figure 4.28: 3D potential profile throughout 'atomistic' N-I-N diode as simulated in chapter 3. Resolved local increase in potential is seen around the marked donors. Potential measured in Volts

The electron concentration illustrated in figure 4.30 closely follows the doping concentration within the highly doped regions. In the same figure, the electron concentration distribution from the atomistic simulation is also compared with that obtained from the continuous simulation. It can be seen that the mean of the electron concentration profile obtained from the atomistic simulation closely follows the electron concentration from the continuous simulation. This builds further confidence in the electron dynamics within the atomistic simulation. Some deviation is observed between 400nm and 450nm and is associated with local reduction in donor density compared with the continuous case at that point.

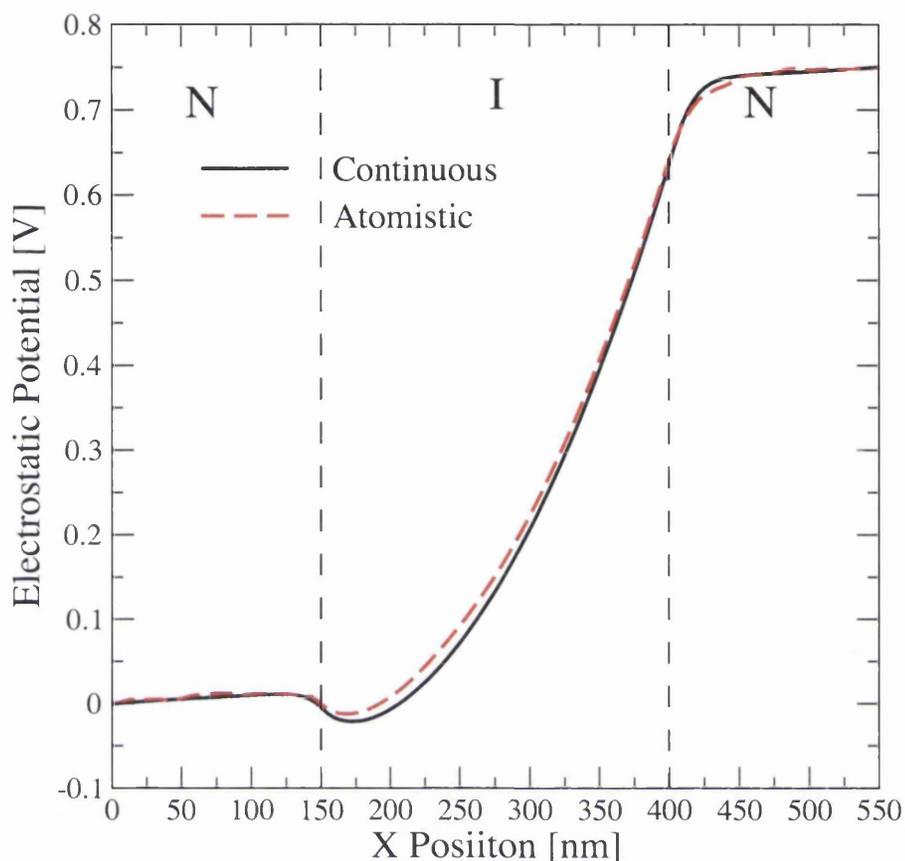


Figure 4.29: Good agreement between electrostatic potentials resulting from the atomistic and continuous simulations are observed and reflect the validity of the method.

4.7.2 'atomistic' electron energy and velocity

Both distributions of the average electron energy and velocity along the length of the atomistic device, plotted in figures 4.31 and 4.32 respectively, are consistent with the continuous simulation. Small discrepancies observed in the atomistic energy profile may be attributable to noise given the limited number of electrons contributing to the statistics. However, the noticeable peak in energy around 250 nm is associated with a close grouping of two donors within the intrinsic region which is in turn seen as a peak in the velocity at the same location. The two donors responsible are highlighted in figure 4.33 which

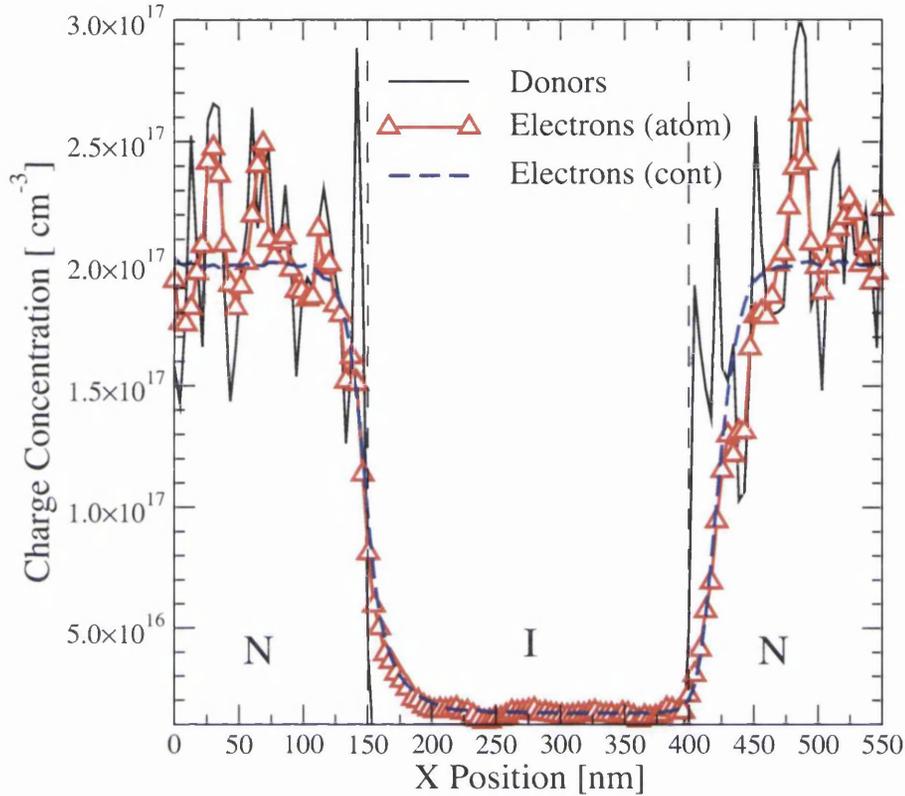


Figure 4.30: Atomistic (atom) electron concentration compared with the atomistic donor concentration and electron concentration from the continuous (cont) simulation. The electron concentration follows the doping with a mean close to the continuous concentration.

shows the donor locations in the device along the x -axis. This variation in energy and velocity adds to the random dopant effects already observed in the potential variation in the depletion regions.

4.8 Conclusion

Within this chapter, the limitation of ionized impurity scattering rates and the continuous description of doping in the simulation of ultra-small devices was discussed. An *ab initio* method that recovers the scattering from a random arrangement of discrete charges,

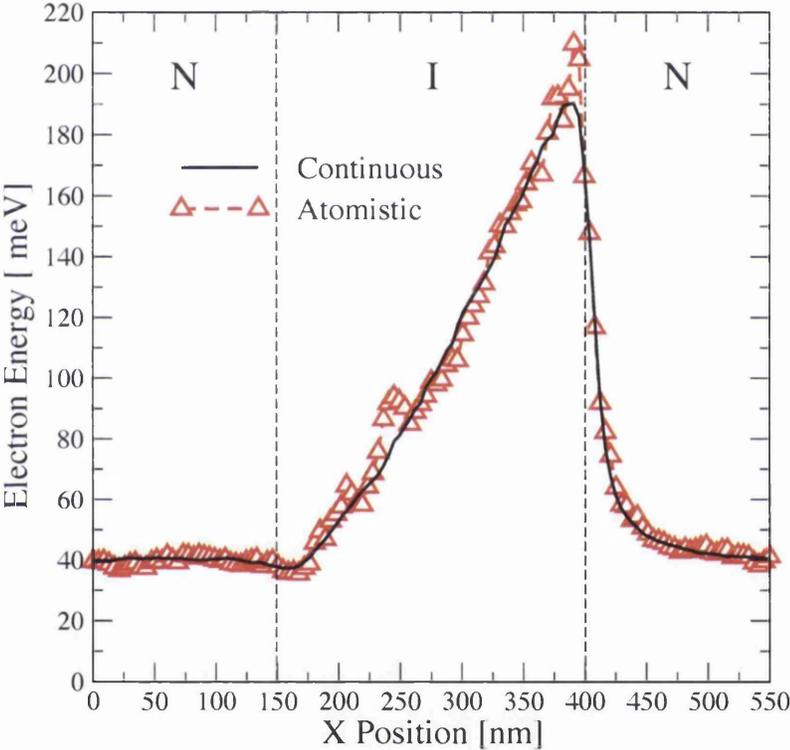


Figure 4.31: Electron energy profile along the device for both atomistic and continuous simulations. Good agreement is seen, with variation observed in the atomistic simulation.

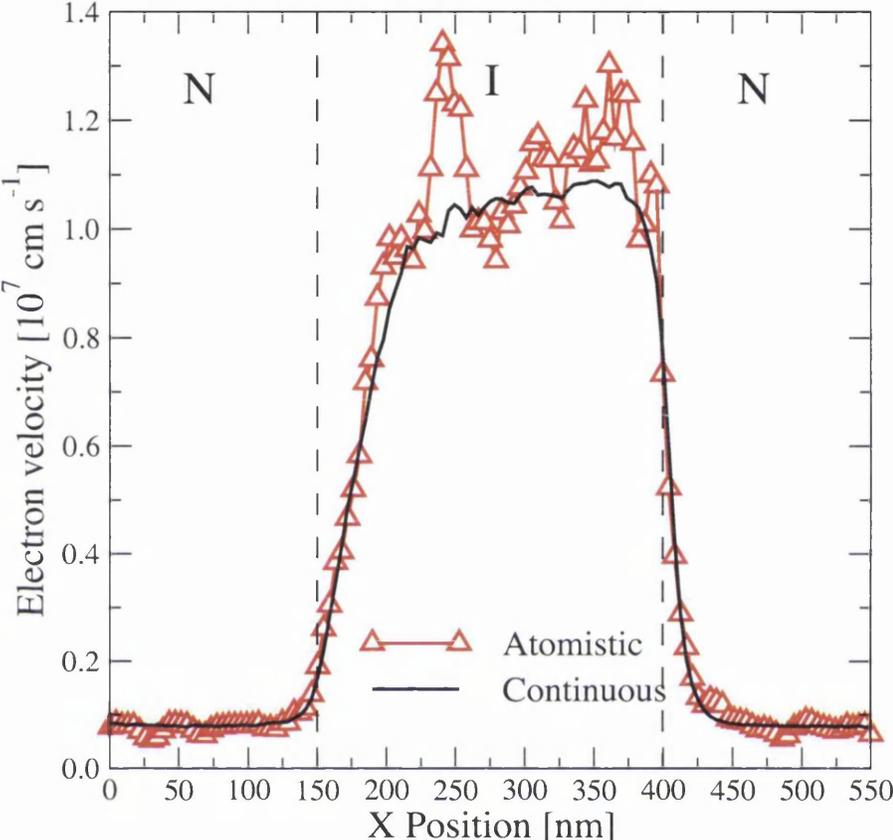


Figure 4.32: Electron velocity profile along the device for both atomistic and continuous simulations. Again good agreement is seen, with higher velocity in the atomistic region.

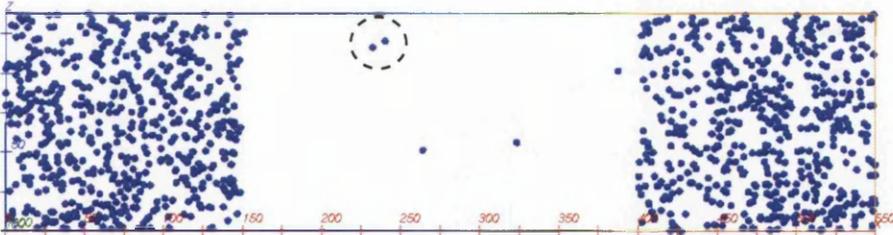


Figure 4.33: Donor positions as viewed along the y-axis. x runs from left to right as indicated on the scale. The two closely spaced donors responsible for the local increase in the energy and velocity profiles (figs. 4.31 and 4.32) are highlighted.

both fixed and mobile, was carefully developed. The integration method used was evaluated under simple test conditions to ensure energy conservation and the reproduction of Coulomb scattering. In order to achieve energy conservation within the limitations of the chosen integration approach, it was necessary to limit the short range Coulomb interaction. Various form of short range interactions were evaluated in respect of both energy conservation and the ability to recover Coulomb scattering. All of them limit the magnitude of large angle scattering, which is justifiable only due to the small probability of such large angle scattering. The application of the short-range correction to the mesh force in order to accurately describe interactions was then discussed and the approach was validated through the reproduction of the low field bulk electron concentration dependent mobility in silicon. The validation was extended to the atomistic simulation and comparison of the N-I-N device simulated using continuous doping in Chapter 3.

Chapter 5

Mobility Variations

5.1 Introduction

A method capable of deterministically and accurately representing the scattering of electrons from a random charge configuration in EMC simulations has so far been discussed. Within the model, Coulomb scattering is incorporated through the trajectories of self consistently propagated electrons in the field associated with the unique charge arrangement. Verification of the model was presented in the last chapter by the reproduction of the doping concentration dependence of the low field bulk mobility of electrons in silicon and by comparing an atomistic simulation of an N-I-N diode to a previous continuous simulation. The *ab initio* bulk mobility simulations were performed using a large enough number of donors in order to approximate a self averaging ensemble in a regime customarily described using ionized impurity scattering rates. In the simulation of ultra-small next generation MOSFETs, there will be only a few tens of dopant atoms present on average in the active region of the device. This represents a system that is not self averaging, prohibiting the use of bulk impurity scattering rates and introducing transport variations from device to device. Therefore, in this chapter the *ab initio* method is applied to the

study of the impact of the doping atomicity in small devices.

Ab initio implementations of Coulomb scattering in transport EMC simulations have recently been reported and used to study a variety of topics. The effect of the long and short range Coulomb interactions upon the thermalization of electrons within the channel have been reported for a short channel MOSFET [80]. The inclusion of the short-range Coulomb interaction led to a significant reduction in electron velocity at the end of the channel due to the electrostatic interaction with the dense, thermal electrons within the drain. Simulation with the *ab initio* Coulomb scattering approach has also been employed in studying current and threshold voltage variations associated with random dopants in a small set of MOSFETs [81]. The work also shows the correlation between threshold voltage shift and dopant position close to the semiconductor/insulator interface. This highlights the extent of the statistical simulations carried out using this method. An alternative technique based upon treating the long range Coulomb interaction through the mesh resolved force and replacing the short-range force correction with an effective scattering rate has been used to reproduce bulk concentration dependent mobility for electrons [142], and to study current variations in small MOSFETs in a deterministic, illustrative way [143, 144].

Despite the above references, there has been no *ab initio* simulations of large statistical sets of devices, like those routinely performed within the Drift-Diffusion (DD) approximation [24, 56, 57], characterising the magnitude of random dopant induced device parameter variations. The lack of statistical data results from the prohibitively long simulation times as compared with other methods, allowing only for simulation of the illustrative examples.

Computational efficiency still remains an important factor in the competitive device design processes and to this end DD simulations are commonly used. The implementation of simple corrections in order to account for basic, yet important, quantum mechanical ef-

fects are often included within DD simulations in an attempt to extend the range of validity when simulating small devices [61–63] and atomicity [59,60]. The question regarding the accuracy of atomistic simulations using DD as compared with the more accurate transport models provided by MC simulation has, however, remained unanswered.

In this chapter, the developed MC simulation methodology with *ab initio* Coulomb scattering is used to study mobility variations due to discrete charges in small volume devices in comparison with DD simulations that only capture associated electrostatic variations. Such mobility variations, if significant, are expected to play an important role in intrinsic parameter variations since there still remains a strong correlation between the mobility and drive current in nano-scale devices [145]. At the beginning of this chapter in section 5.2, a new charge assignment scheme for continuous doping is proposed in order to consistently represent combinations of both continuous and atomistic doping within the same simulation domain. The impact of a single trapped charge in a MOSFET is investigated in section 5.3. Finally a statistical set of devices with fully atomistically doped channels is simulated in section 5.4.

5.2 Continuous-Cloud-In-Cell Charge Assignment

In the structures simulated in this chapter it is necessary to combine both continuous and atomistic doping. Indeed, combining the two doping descriptions within the same simulation may be advantageous in general. In such combination however, care needs to be taken to ensure the mesh assigned doping charge is consistently treated. Otherwise a discrepancy arises in combining the charge assigned to represent the continuously doped mesh cells and the Cloud-In-Cell approach used to represent the individual dopant atoms. This discrepancy results from the different manner in which the two doping types are attributed to the mesh nodes in the discretization of the Poisson equation. The same incon-

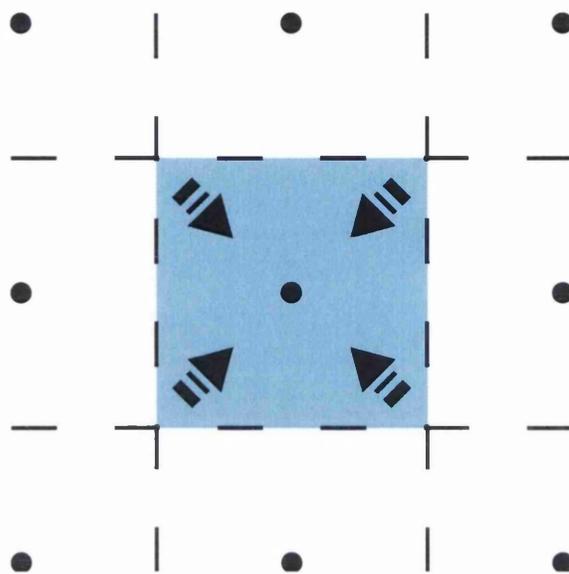


Figure 5.1: Assignment of charge from the uniform region bound by the mesh cell to the associated node is essentially the NGP charge assignment scheme, when considering the node centred mesh.

sistency also applies to the assigned electron charge within a continuously doped region and is therefore important within traditional MC simulations. Inconsistencies in the mesh assignment lead to unwanted variation in charge density and therefore in electrostatic potential. Such artefacts are detrimental to accurate and physically meaningful simulations and should be removed. Therefore, in the simulation of small devices/volumes on fine meshes, it is necessary to use a consistent charge assignment scheme for all discrete charges and continuous charge densities. Such a treatment is dependent upon the choice of mesh description and is discussed next for the two choices discussed in chapter 2.

5.2.1 Node Centred Mesh

Traditionally, the contribution of charge from the continuous doping assigned to any mesh point is given by the amount of charge contained within the cell volume associated with

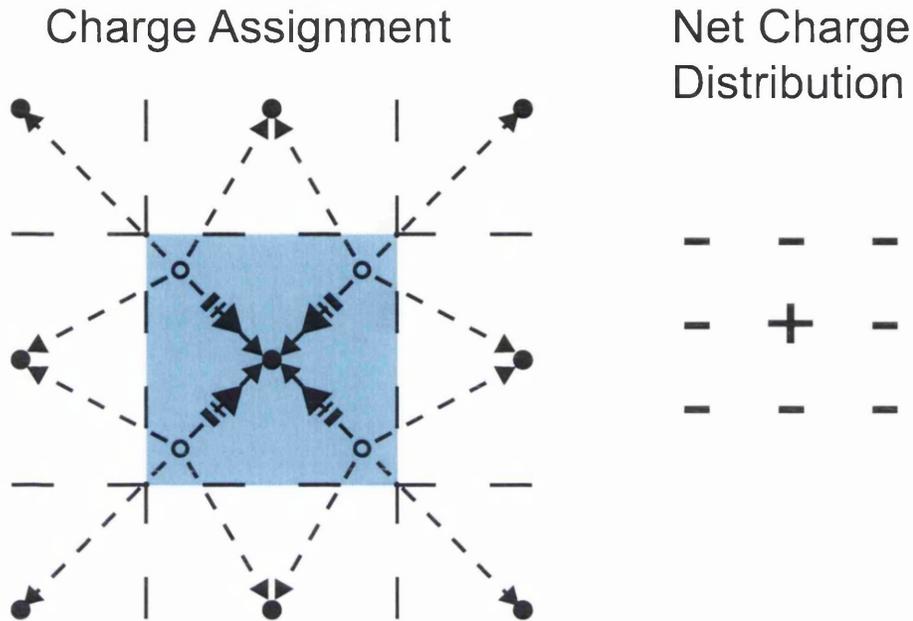


Figure 5.2: Though the positions of electrons with respect to the fixed background charge may satisfy charge neutrality, the difference in the charge assignment to the mesh guarantees that neutrality can not occur at the edge of doping regions. Thus irrepressible fluctuations within the electrostatic potential occur.

that mesh point. This is simply determined as the product of the specified charge density and cell volume. For the first choice of mesh described in chapter 3 and advocated by Hockney and Eastwood, mesh points are centred within the cell. Cell boundaries are parallel to mesh lines but situated half way between the mesh points. In this case this continuous charge assignment can be regarded as equivalent to the Nearest-Grid-Point (NGP) scheme applied to the total charge within the cell volume as illustrated for a 2D mesh in figure 5.1. This scheme conflicts with the Cloud-In-Cell (CIC) scheme used to assign the charge associated with both electrons and discrete donors, and this conflict is in turn responsible for the extended electron charge at abrupt heterojunctions, previously mentioned as a potential problem of the CIC scheme (chapter 3).

The problem arises because, within schemes of higher order than NGP, charge associ-

ated with a particle is assigned over surrounding mesh points, hence neighbouring cells, while continuous doping charge is assigned to just one cell. As a result of this conflict, illustrated in figure 5.2, electron charge assigned via the CIC scheme can never accurately balance donor charge at the edge of a continuously doped region. This may lead to significant unwanted artefacts within the mesh based potential.

Similarly, though more importantly in this case, the same confliction results in a poor description of the donor charge density when the doping changes from a continuous medium to discrete charges. At the boundary between continuous and discrete doping, CIC assignment from the discrete charges overlaps with the charge assigned from the continuous medium, which ends abruptly due to the NGP-like assignment. As illustrated in figure 5.3, this results at the boundary between the two regions in an overestimation of the charge density in the continuous medium, followed by an underestimation in the discrete medium. These unwanted artefacts are mesh dependent leading to larger fluctuations in the potential associated with smaller mesh spacings.

It is therefore desirable to modify the charge assignment in the continuous charge medium in order to reproduce the CIC scheme. Such modification should simply result in smoothing of the continuous charge over an extra mesh point and can be achieved by using a smoothing kernel which is propagated over the initial, NGP assigned, charge regions. The application of the smoothing kernel re-apportions to a central mesh point the contribution of charge within surrounding cells that would otherwise have been assigned to it using the CIC scheme. The elements of the smoothing kernel are then simply the fraction of continuous charge contained within the central mesh cell, and those cells surrounding it, that need to be re-apportioned to the central mesh point.

The elements of the smoothing kernel can be determined by considering that the continuous charge consists of a regular array of uniform sub-charges with charge equal to $Qdx dy dz/V$, where Q is the continuous charge within the whole cell and dx, dy, dz are

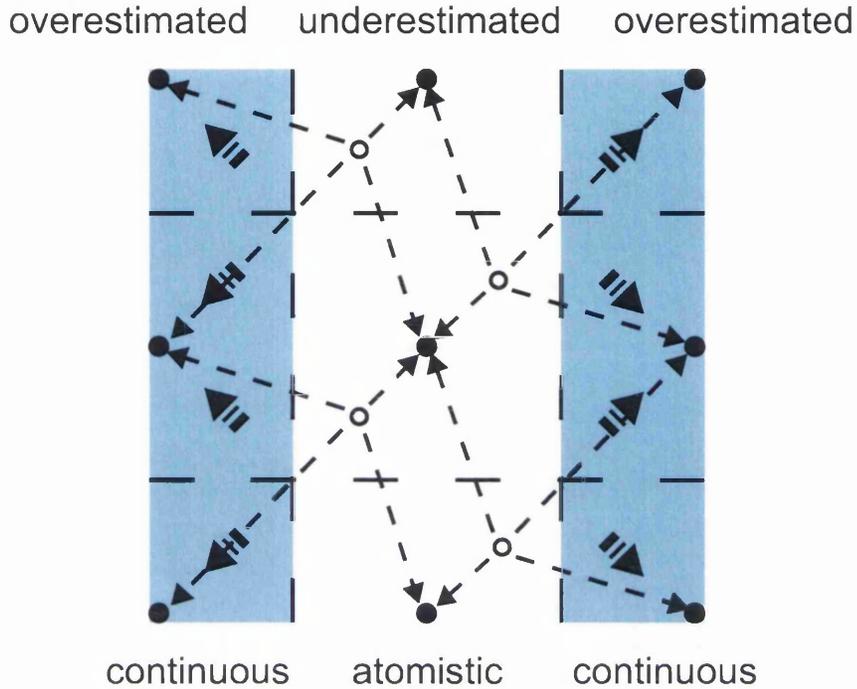


Figure 5.3: Charge assigned from the discrete dopants overlaps with that assigned to represent the continuous doping. Overlap from the continuous doping to the discrete doping does not occur and leads to over/under estimation of the dopant charge density.

dimensions of each sub-division within the cell. Each sub-charge is then assigned not via the NGP scheme but via the CIC scheme as illustrated in figure 5.4. Within this scheme, charge assigned to any mesh point follows from the presence of a sub-charge within one mesh spacing of the point in all directions. The elements of the smoothing kernel are obtained by taking the integral of the CIC assignment function, applied now to a regular array of infinitesimal sub-charges, over the areas associated with each contributing cell given by

$$\frac{Q}{V} \int_{x_0}^{x_1} (1 - |x|) dx \int_{y_0}^{y_1} (1 - |y|) dy \int_{z_0}^{z_1} (1 - |z|) dz \quad (5.1)$$

where $x_0, x_1, y_0, y_1, z_0, z_1$ are the x, y and z limits of the contributing cell in the CIC assignment region with respect to the central mesh point. In a 3D mesh there are twenty seven

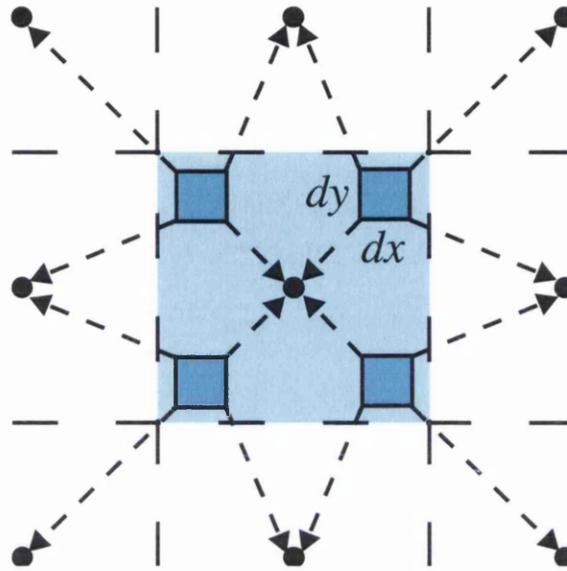


Figure 5.4: Continuously doped charge throughout the mesh cell is assigned via the CIC manner, consistently with the discrete charge.

contributing regions to consider that, due to the symmetry of the assignment scheme, reduce to four distinguishable regions highlighted in figure 5.5. The corresponding limits are tabulated in table 5.1. Region 1 is associated with the central mesh cell itself, where integration in this region is performed over one eighth of the total cell volume. Of the neighbouring cells, there are 3 distinguishable types. Six cells that are directly adjacent to the central cell and share a common cubic face define region 2. The integration over region 2 is over a quarter of the total volume associated with these neighbouring cells. Region 3 consists the eight neighbouring cells sharing a common edge and the integration over this region covers half the total volume associated with these cells. Finally, region 4 is the volume associated with the eight neighbouring cells sharing a common corner and the integration is carried out over this entire region. Carrying out the integration defined by equation 5.1 over these regions, and multiplying by the fraction of the total contributing volume they represent, yields the elements of the smoothing kernel which are compactly

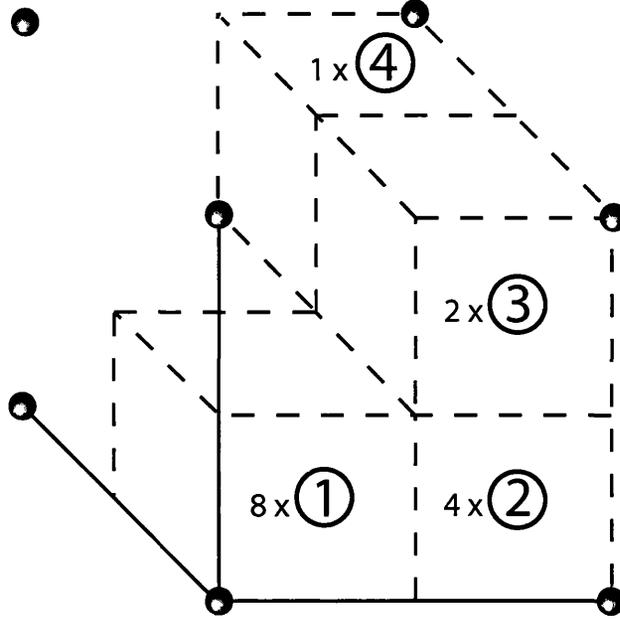


Figure 5.5: Unique regions over which integration around a node, tabulated in table 5.1, is performed.

written for arbitrary dimensions N as

$$\rho_{x,y,z} = \left(\frac{3}{4}\right)^N \sum_{i,j,k=-1,0,1} f(i) f(j) f(k) \rho_{x+i,y+j,z+k} \quad (5.2)$$

where

$$f(a) = \begin{cases} \frac{1}{6}, & a = \pm 1 \\ 1, & a = 0 \end{cases} \quad (5.3)$$

and the number of such functions $f(a)$ should equal the number of dimensions. Application of this smoothing kernel to an NGP assigned continuously doped region is termed the Continuous-Cloud-In-Cell (CCIC) charge assignment process. In 3D the kernel is a $3 \times 3 \times 3$ matrix that is applied to a central mesh point and the 26 surrounding mesh points. At the edge of the discretisation mesh, where surrounding points are absent, the smoothing kernel is passed over reflections of the mesh points about the boundaries in

Region	x_0	x_1	y_0	y_1	z_0	z_1
1	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$
2	$\frac{1}{2}$	1	0	$\frac{1}{2}$	0	$\frac{1}{2}$
3	$\frac{1}{2}$	1	$\frac{1}{2}$	1	0	$\frac{1}{2}$
4	$\frac{1}{2}$	1	$\frac{1}{2}$	1	$\frac{1}{2}$	1

Table 5.1: Unique regions and limits of integration associated with cells contributing to the assigned charge at a node. Regions refer to those illustrated in figure 5.5

order to ensure conservation of charge. As this charge assignment is consistent with the CIC assignment from a point charge, application of this smoothing guarantees an accurate transition between all doping regimes and ensures that the electron charge density can balance the donor charge density at any heterojunction.

5.2.2 Node Aligned Mesh

For the mesh used in this work, the mesh boundaries are aligned along lines connecting mesh points rather than lines half way between mesh points. Any cell is now defined within these boundaries and this changes the description of the material on the mesh. Unlike the previous case, where assigning the continuous doping charge to a grid node could be thought of as assignment by the NGP scheme, the charge within a cell is now assigned entirely to the lowest integer grid point in each dimension that bounds the cell. This is illustrated in figure 5.6 (cf. figure 5.1). Thus, although the cell boundaries include multiple mesh points, the properties of the cell are still associated with just one in this context. CIC assignment of a point charge still implies assignment to the neighbouring cells in each direction. Thus the problem of assigning charge over an abrupt heterojunction within the continuous doping regime is still present.

Similarly to the previous case, the junction between a continuous doping region and

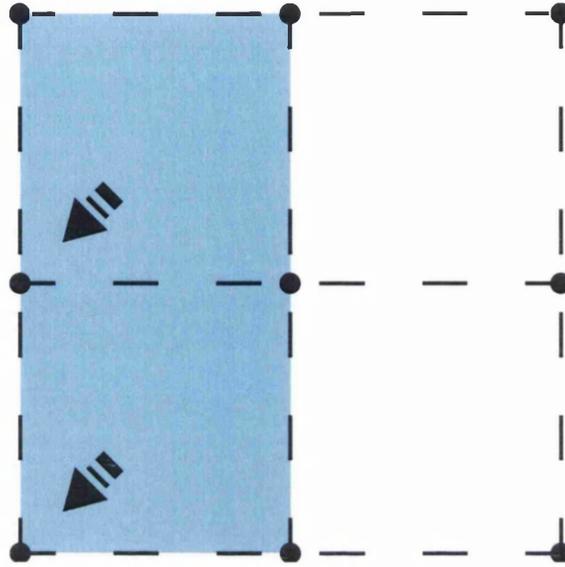


Figure 5.6: Assignment of continuous charge within the cell is entirely to the lowest integer node in all dimensions within the cell. This differs from figure 5.1

an atomistic doping region requires attention. Combining a continuously doped region with the CIC scheme for atomistic doping assignment results in two types of boundary artefacts being imposed due to the lack of symmetry in the continuous charge assignment as is illustrated in figure 5.7. If the continuous region precedes the atomistic region on the mesh, the charge density is under-represented at the mesh point defining the boundary. Alternatively, if the continuous region precedes an atomistically doped region, the charge density at the boundary is overestimated.

By properly accounting for the assignment of the continuous charge in the same manner as before (see figure 5.8) such problems are avoided. As the cell description is different, the bounds of integration over the regions of contributing charge in equation 5.1 have to be modified. From the mesh construction and the symmetry of the CIC assignment, each of the eight cells that in 3D borders a common mesh point contributes the same fractional amount of the charge to that mesh point. Therefore, instead of four unique con-

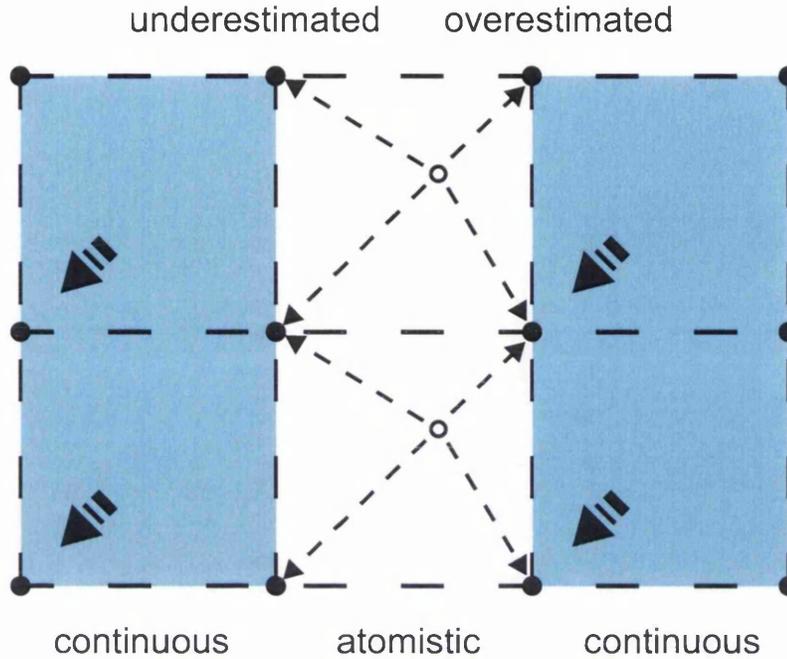


Figure 5.7: Charge is under represented at the transition from continuous to atomistic doping due to the difference in the charge assignment schemes.

tributing regions as is the case for the previous mesh, there is only one. The new region of integration, illustrated in figure 5.8, is over an entire cell adjacent to the mesh with integration limits given in table 5.2. Applying these limits yields a simple smoothing function that similarly reproduces the CIC assignment.

$$\rho_{x,y,z} = \left(\frac{1}{2}\right)^N \sum_{i,j,k=-1,0} \rho_{x+i,y+j,z+k} \quad (5.4)$$

Again N is the number of dimensions and the subscripts labelling the dimensions coincide accordingly. In contrast to equation 5.2, used for the node centred mesh description, in 3D equation 5.4 represents a $2 \times 2 \times 2$ matrix of identical elements that is applied only over the principal and the preceding mesh points. At the edge of the discretisation mesh the kernel is passed over copies of the available nodes about the edge to ensure charge

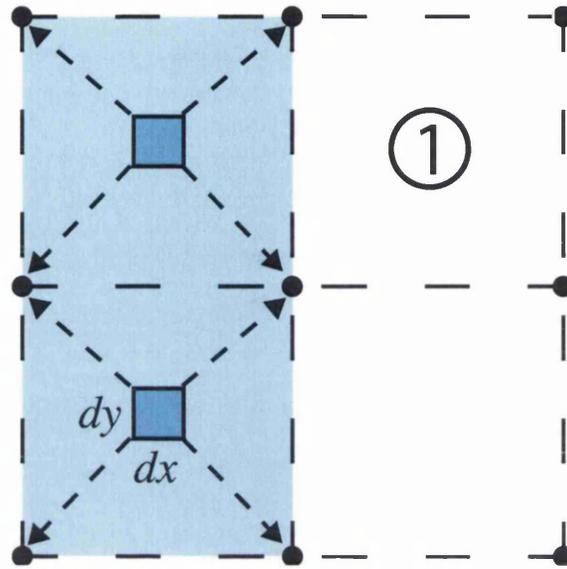


Figure 5.8: Continuous charge within a cell is evenly distributed among the surrounding nodes (cf. figure 5.4). The fraction of charge assigned to the central mesh point from each surrounding cell is determined by integrating over 1 unique region.

Region	x_0	x_1	y_0	y_1	z_0	z_1
1	0	1	0	1	0	1

Table 5.2: Limits of integration over the unique region for the evaluation of the fraction of charge contained within a cell to be associated with a mesh point in order to replicate CIC assignment.

conservation. This form of the CCIC smoothing evenly distributes the charge contained within a cell to the surrounding nodes (see figure 5.8) and is specific to the choice of mesh and material property identification adopted within this work. It has been applied consistently throughout the next simulations.

5.2.3 Application of CCIC Charge Assignment

The effect of applying the CCIC smoothing on the doping configuration illustrated in figure 5.7, the transition between continuous to atomistic and back to continuous dop-

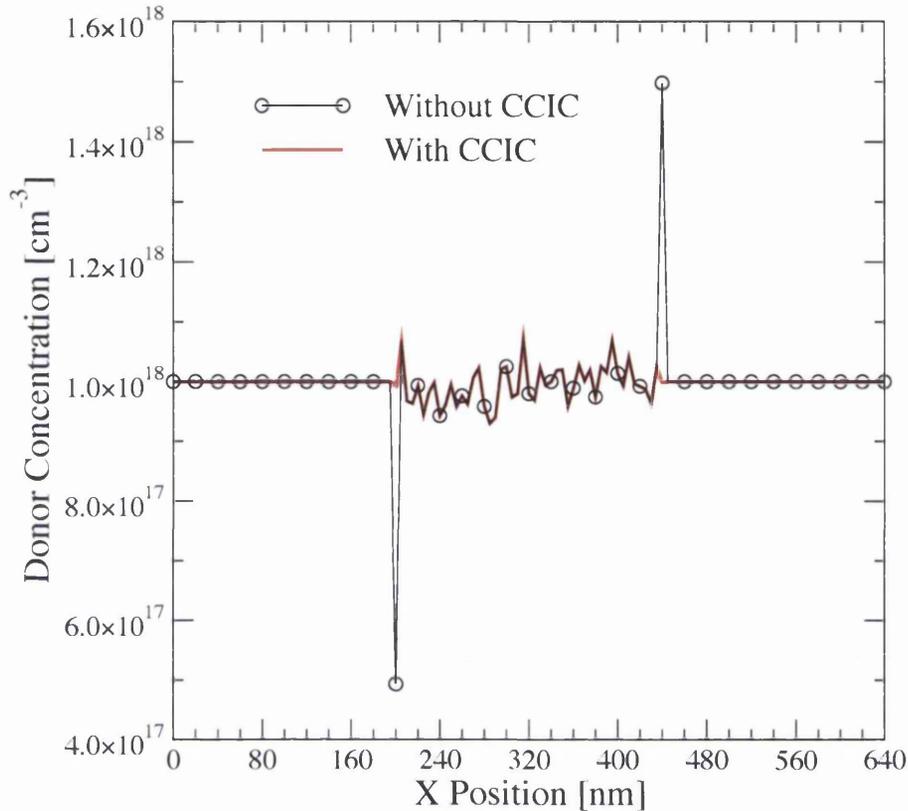


Figure 5.9: The donor concentration profile using a mixture of continuous and atomistic doping, with and without applying the CCIC charge assignment function to the continuous donor charge.

ing, is now shown. A uniform mesh with mesh spacing $H_x = H_y = H_z = 5 \text{ nm}$ was constructed consisting of $129 \times 65 \times 65$ mesh points. For $x < 200 \text{ nm}$ and $x \geq 440 \text{ nm}$ the volume was doped with a continuous donor concentration equal to $1 \times 10^{18} \text{ cm}^{-3}$. For $200 \text{ nm} \leq x < 440 \text{ nm}$ the doping was considered as atomistic at the same concentration. The doping concentration profile for this structure is shown in figure 5.9 with and without the application of the CCIC assignment to the continuous donor charge. Without the smoothing, the transition from continuous to atomistic doping results in a significantly undervalued donor charge concentration, while the subsequent transition from atomistic to continuous shows significant overestimation. The underestimated concentra-

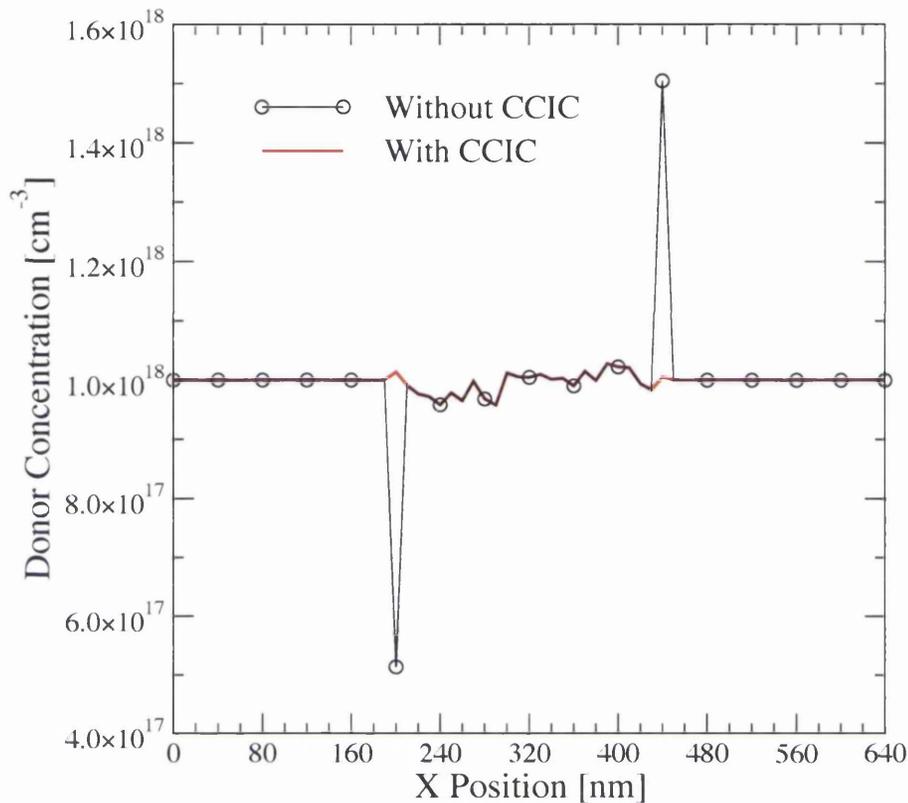


Figure 5.10: Same donor configuration as in figure 5.9 but assigned to a more coarse mesh. The same inconsistency is seen at the transition from continuous to atomistic doping. Application of the CCIC function again removes this.

tion is $5 \times 10^{17} \text{ cm}^{-3}$ while the overestimated concentration is $1.5 \times 10^{18} \text{ cm}^{-3}$, representing $\pm 50\%$ of the required value. This is as expected from the difference in the assignment schemes at the two boundaries, and upon applying the CCIC assignment the transitions between the regions are accurately accounted for.

Figure 5.10 shows the same donor configuration, but this time on a mesh twice as coarse with $65 \times 33 \times 33$ mesh points and a resulting uniform mesh spacing of 10 nm . The same consistent error of $\pm 50\%$, is seen at the transition between the continuous and atomistic doping regions, introduced by the assignment of the charge contained within each cell. However, this variation in charge density results in a less significant variation in

the potential and field in this region due to the larger mesh spacing. In this way the artefact in the charge assignment introduces mesh dependent variations in potential. Again, the application of the CCIC assignment function results in an accurate transition between the two regimes.

In conclusion, the proposed smoothing of the continuously assigned doping charge density maintains consistency in the assigned charge densities associated with either dopant atoms or electrons and continuous charges regions. A smoothing kernel has been evaluated which transforms the NGP assigned continuous charge density to match the CIC assignment scheme used in the case of discrete charges. The result of the new Continuous-Cloud-In-Cell charge assignment improves the doping charge density description in transition from continuous to atomistic doping and removes mesh based artefacts in the potential associated with inconsistently assigned charges.

5.3 Single Trapped Charge

In this section, the developed 3D MC simulator featuring *ab initio* Coulomb scattering is used to evaluate the impact of transport variation associated with scattering from a trapped charge on the magnitude of the corresponding drain current reduction for a series of well scaled *n*-channel MOSFETs. By comparison with Drift Diffusion (DD) simulations that only capture the electrostatic effects associated with the trapped charge, the contribution to the current reduction due to scattering is demonstrated. Simulations are performed at low drain bias for MOSFETs with channel lengths of 30 nm, 20 nm and 10 nm. Compared to DD, MC simulation shows significant additional reduction in drain current. The reduction in current due to scattering increases with the gate voltage towards strong inversion conditions.

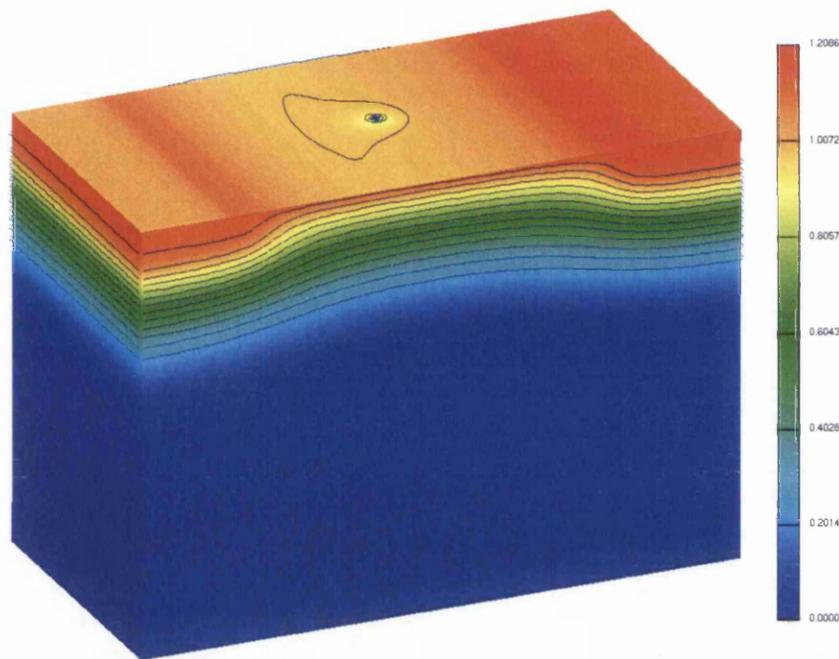


Figure 5.11: DD solution of the potential throughout the simple 30 nm MOSFET. The location of the trapped charge is clearly discernable as the region of lower potential in the centre of the channel.

5.3.1 Introduction

Mass production MOSFETs will reach nanometer (sub-10 nm) dimensions near the end of the International Roadmap for Semiconductors (ITRS) [1]. Trapping of a single carrier charge in defect states near the Si/gate dielectric interface and the related local modulation in carrier density and/or mobility [73, 74] will have a profound effect on the drain and gate current in such devices [64]. Corresponding Random Telegraph Signals (RTS) with amplitudes as large as 60% have been already observed experimentally [67]. This problem will be exacerbated by the higher defect density in high- κ materials, which are expected to replace SiO₂ in the gate stack somewhere between the 65 nm and the 45 nm technology nodes [146]. Current fluctuations on such a scale will become a serious issue, not only as a source of excessive low frequency (LF) noise in analogue and mixed-mode

L (nm)	T_{ox} (nm)	N (cm^{-3})
30	1.00	5×10^{18}
20	0.75	1×10^{19}
10	0.50	3×10^{19}

Table 5.3: Length scales and parameters of the simple MOSFET structure used within this simulation.

circuits [147, 148], but also in dynamic and SRAM memories [149] and other digital applications. Depending on the device geometry, a single or few discrete charges trapped in hot carrier or radiation created defect states will be sufficient to cause significant performance degradation in decananometer and nanometer scale MOSFETs [150, 151].

Until now the 3D simulation studies of the effects associated with trapping of a single charge in the channel of decanano MOSFETs, and the corresponding RTS amplitudes, were carried out under the DD approximation [65]. Such simulations only capture the electrostatic creation of an inversion layer exclusion region around the trapped charge which reduces the overall current flowing through the device. It is still debatable in the literature as to whether the electrostatic reduction in the carrier density or increased scattering in the channel dominates the reduction in current in response to the charge trapping and in determining the magnitude of the RTS amplitudes [72–74].

The developed MC simulator, explicitly including Coulomb scattering through the real space trajectories of carriers, is here applied to study the transport effects associated with scattering from the trapped electron. The MC simulation incorporates the same electrostatic effects as DD but, importantly, the additional Coulomb scattering from the trapped charge is also incorporated. By comparing the MC and DD simulations, it is possible to isolate the specific contribution of the additional Coulomb scattering within the current reduction.

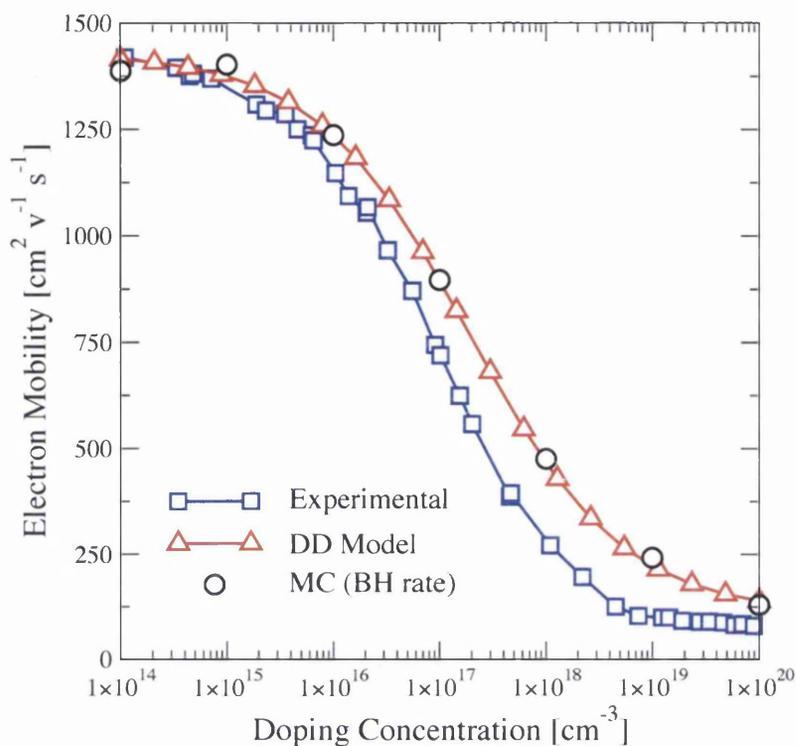


Figure 5.12: Caughey-Thomas mobility model fit to the bulk MC result obtained using the Brooks-Herring scattering rate. This serves to allow comparison of DD and MC results.

5.3.2 Simulation methodology

In order to study the impact of scattering on the current reduction associated with charge trapping, three *n*-channel MOSFETs with square channels measuring 30 nm, 20 nm and 10 nm have been simulated using both 3D DD simulations and MC simulations featuring *ab initio* Coulomb scattering. Each device has a simple structure and is well scaled to achieve good electrostatic integrity. The basic device parameters are given in table 5.3. A single negative charge is placed in the centre of the channel at the Si/SiO₂ interface and represents a trapped electron. In this position the electrostatic impact on the current flow within the channel at low drain voltage is most pronounced [65]. Continuous doping is otherwise employed within the channel and in the source/drain regions. Only the

Coulomb scattering from the trapped electron is *ab initio* included in the Monte Carlo simulations, otherwise conventional ionised impurity scattering rates are used throughout the device to represent the scattering associated with the continuous doping. Within the DD simulations, analytical concentration and lateral-field dependent mobility models are used. The parameters of these models are chosen to best fit the concentration and field dependent mobilities obtained through bulk MC simulation, rather than fitting to experimentally observed values. In this way DD and MC simulations can be fairly compared. Comparison of the analytical mobility models with bulk MC results for concentration dependent and field dependent mobility are shown in figures 5.12 and 5.13 respectively. No perpendicular-field dependence of the mobility is used in the DD simulations as this implies the action of surface roughness scattering which is not included within the MC model.

The simulations are carried out at a low drain bias of $50mV$ for all devices in order to minimize non-equilibrium transport and maintain a fair comparison between DD and MC results. In this case there is little difference in the resulting self-consistent potential and electron concentration between the two simulation techniques. That this is the case is seen by comparing the electron concentration resulting from DD and self consistent MC simulation in figures 5.14 and 5.15. Therefore, in order to reduce simulation times, MC simulations are performed in the frozen-field approximation where the field used is obtained from a prior self-consistent DD solution. In all simulations, Poisson's equation is solved on a uniform mesh with a mesh spacing of $0.5nm$. Since the MC simulations are not self-consistent, the Coulomb potential of the individual carriers remains unresolved and the carrier-carrier interactions are excluded. The frozen-field approximation is further validated by comparing results obtained using self-consistent DD and frozen-field MC simulations. The gate voltage dependence of the drain current for the $30nm$ MOSFET obtained from both methods is plotted in figure 5.16. The agreement over the range of gate

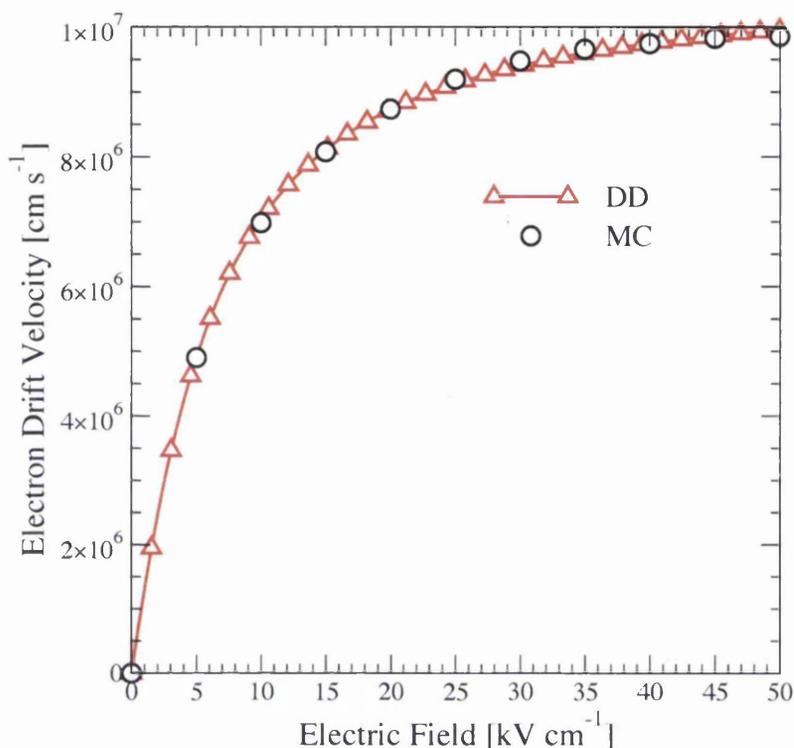


Figure 5.13: Fit of the velocity-field curve to bulk MC simulation, allowing comparison between DD and MC results.

voltages is good with MC simulations giving consistently higher current, overestimating DD by around 20%. Comparisons for the 20nm and 10nm device show similar agreement at high gate voltages but larger disagreement at low gate voltages. This is related to significant carrier heating in these very short devices, even at 50mV drain voltage, which cannot be captured in the DD simulations. From these results it is argued that the frozen-field approximation deliver sufficient accuracy to allow a fair comparison between the DD and the MC simulations, particularly at high gate voltage.

To conduct the study, the drain current is estimated over a range of gate voltages for each device, with and without a single trapped electron present in the centre of the channel. The drain current is estimated using both self-consistent DD simulations and frozen-field MC simulations using the corresponding DD potential solution. For the two

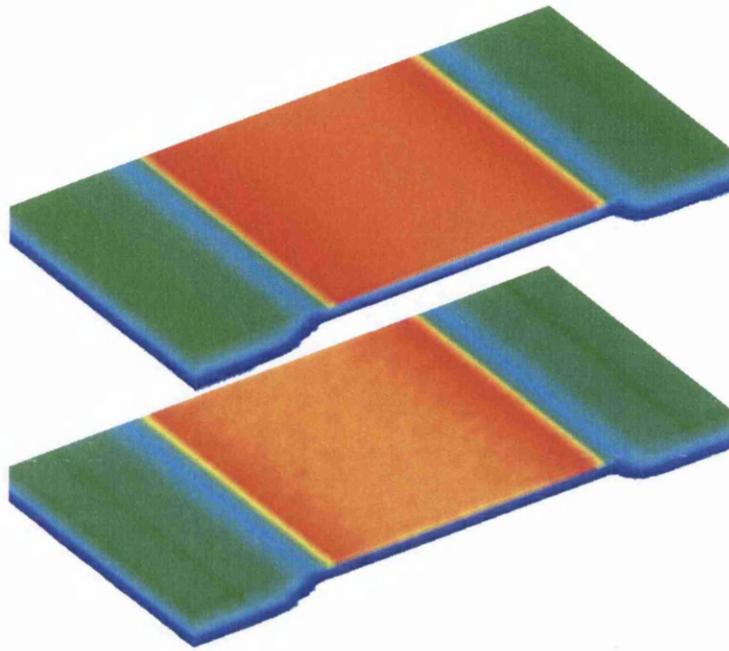


Figure 5.14: Electron concentration obtained from Drift Diffusion (top) and self-consistent Monte Carlo simulation (bottom) within the source, channel and drain of the 30nm device showing good agreement.

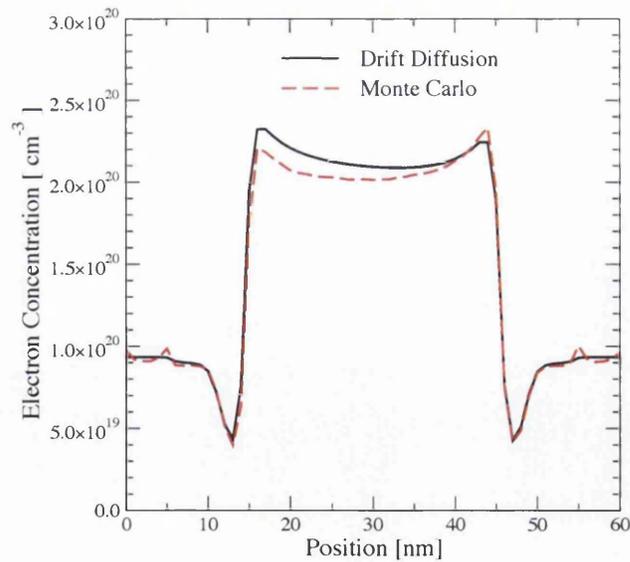


Figure 5.15: Electron concentration profile from DD and self consistent MC simulation using the same data presented in figure 5.14. Good agreement validates the frozen field approximation.

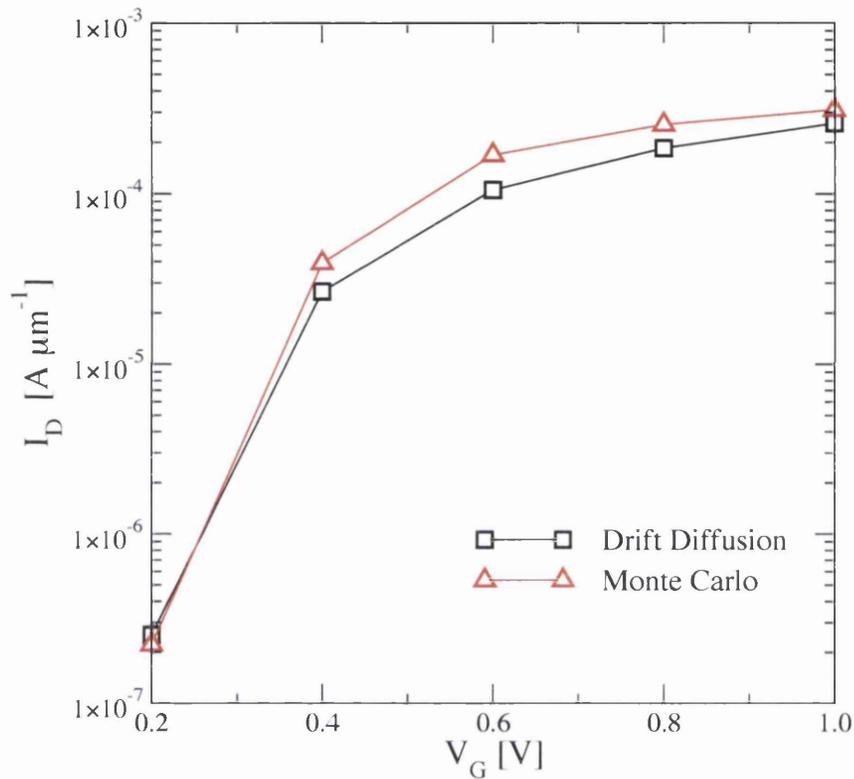


Figure 5.16: Current characteristics for the 30nm device as determined from DD and frozen field MC simulation. The results are comparable, with MC typically giving a larger current, again validating the frozen field approach taken.

simulation techniques, the percentage reduction in current upon including the trapped charge is then evaluated and plotted as a function of the gate voltage. Comparison and interpretation is performed on these results.

5.3.3 Results

5.3.3.1 Percentage Current Reduction in 30nm MOSFET

The self-consistent DD solutions precede the frozen-field MC simulations. As previously stated, the DD simulations only account for the electrostatic reduction in the local carrier density around the trapped charge. The potential from a typical DD solution is shown in

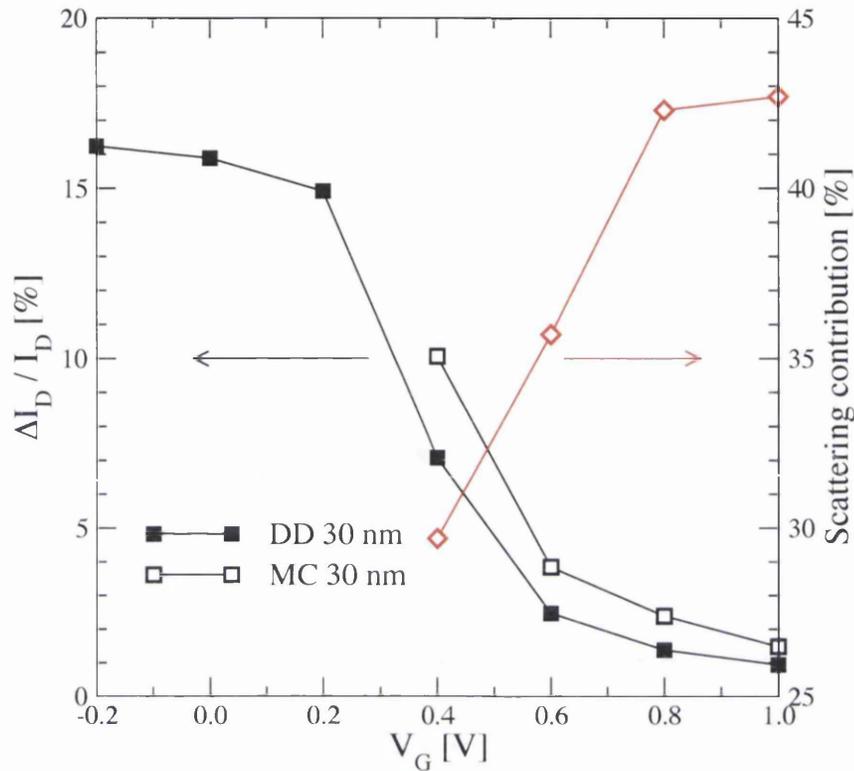


Figure 5.17: Percentage reduction in current upon introducing a trapped charge in the centre of the 30nm MOSFET for a series of applied gate voltages. Results are shown for both MC and DD simulation. Also plotted is the contribution of scattering to the total reduction from MC

figure 5.11 at the start of this section and clearly shows the potential-well in the centre of the channel. This simply increases the resistance of the channel by the exclusion of a circular part of the inversion layer around the trapped electron which leads to the current reduction [65]. The percentage reduction in current as a function of the applied gate bias upon introducing the trapped charge is plotted in figure 5.17.

At low gate voltages, in the subthreshold regime, the inversion layer density is low and cannot efficiently screen the trapped charge. This results in a large charge exclusion region and a correspondingly large reduction in current. At higher gate voltages, in the strong inversion regime, the increased screening from the high density of carriers

in the inversion layer reduces the radius of the exclusion region surrounding the trapped charge. Consequently there is a decrease in the percentage current reduction. The effect of screening is therefore to dramatically reduce the RTS amplitude which, for example, in the 30nm MOSFET drops from 15% in the subthreshold region to less than 1% in strong inversion.

The frozen-field MC simulations follow the DD simulations using, for each device and each bias point, the potential obtained from the DD simulations. Unfortunately the MC approach is very inefficient in the subthreshold regime where the channel carrier density is very low, requiring excessively large simulation times in order to accumulate sufficient statistics to allow an accurate estimation of the current. Therefore, the MC simulations are restricted to the range in gate voltages from 0.4V to 1.0V . Within the MC simulations, similarly to the DD simulations, the inclusion of the trapped electron results in a reduction of the drain current as also plotted in figure 5.17.

5.3.3.2 Separation of Electrostatic & Coulomb Scattering Effects

The same qualitative dependence on gate voltage is observed, with lower percentage reduction in current at higher gate voltages due to greater screening of the trapped charge. However, the percentage reduction is consistently larger compared with the DD results. Assuming the electrostatic dependence is the same in both simulations, the difference between the MC and DD results is due to additional scattering from the screened Coulomb potential of the trapped electron. The difference can be expressed as a percentage of the total current reduction estimated from MC and can be used to evaluate the importance of scattering over electrostatic charge exclusion. This percentage scattering contribution is also plotted in figure 5.17. At $V_G = 0.4\text{V}$, it is seen that the contribution from scattering accounts for 30% of the total current reduction. Here, where screening is weak, the electrostatic exclusion of carriers within the channel is the dominant mechanism in reducing

the current. As the gate voltage rises to $V_G = 1.0V$, and the screening rises with carrier density within the channel, the percentage contribution from scattering rises to around 43 %, highlighting the increased importance of Coulomb scattering. Thus, while the electrostatic influence of the trapped charge determines the response of the current reduction to the applied gate bias, additional scattering increases the current reduction at all biases. The electrostatic influence is dominant in the near threshold regime, and arguably into the sub-threshold, while at high gate voltages, corresponding to strong inversion, scattering from the screened Coulomb potential becomes increasingly dominant and cannot be ignored.

It is unclear as to why the scattering contribution levels off around 43 % at the highest gate voltage. It should be made clear that the statistical error in the current estimation is of the order of 1 – 2 % and lies within the symbols plotted in figure 5.16. However, the simplification of the frozen-field approximation can not be ignored in the presence of no plausible physical explanation.

The component of current density from source to drain at the interface resulting from MC simulation is shown for the cases of $V_G = 0.4V$ and $V_G = 1.0V$ in figures 5.18 and 5.19 respectively, both with and without a trapped charge. In both cases, the trapped charge reduces the current density in the vicinity of the charge with a symmetry reflecting the symmetry of the screened charge potential. The current density resulting from DD simulation is shown for the same gate biases but only including a trapped charge in figure 5.20. The same dependence is seen in these cases, with a larger region of low current density around the charge associated with the larger screening length at low gate voltages.

The current density profile from MC simulation along a line from source to drain through the position of the trapped charge is shown, both with and without the trapped charge present for all gate voltages, in figure 5.21. The *pn* junctions marking the source /channel/ drain boundaries are marked by vertical dashed lines. The local reduction in

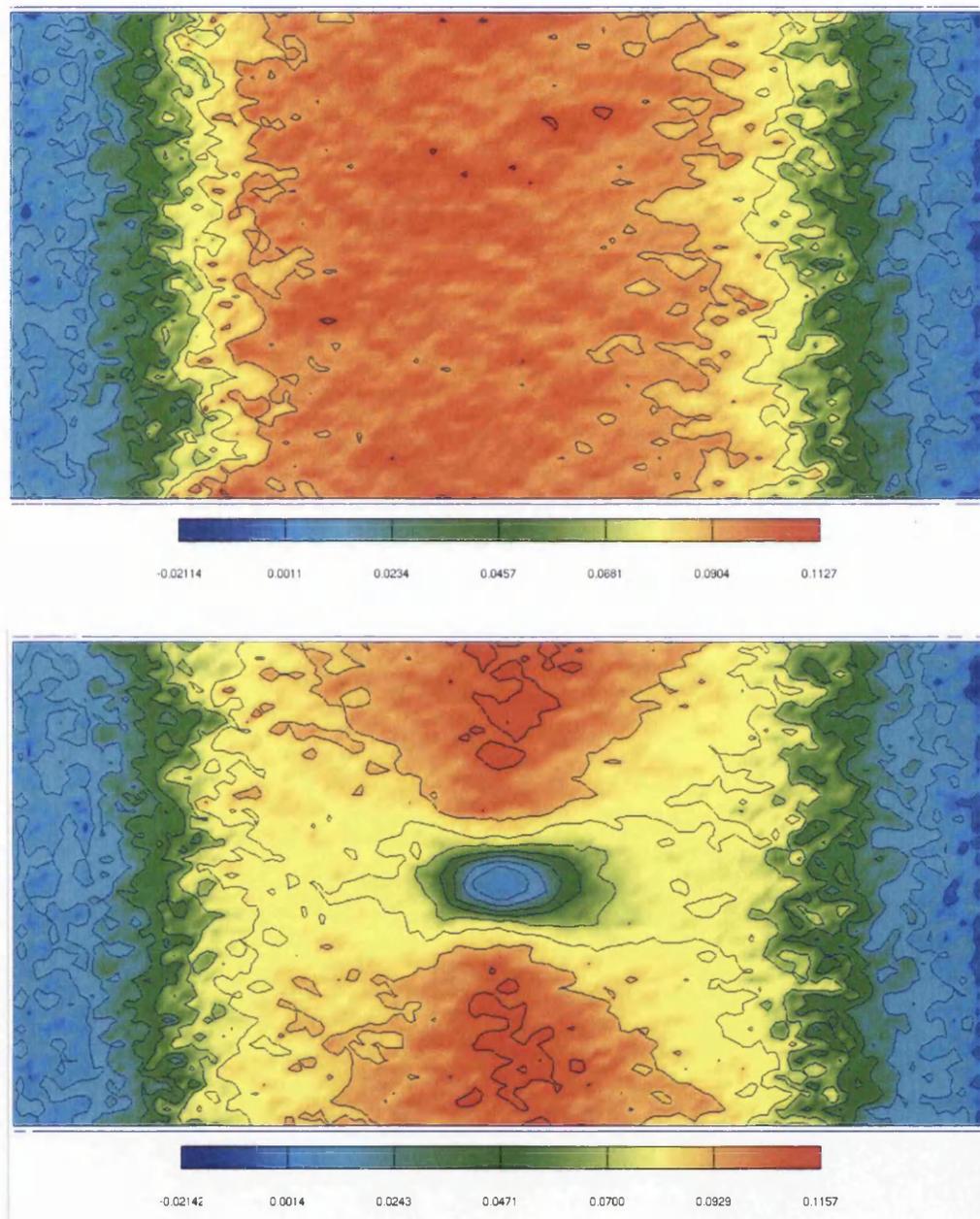


Figure 5.18: x -component of the average electron current density at the interface both omitting (top) and including (bottom) the trapped charge. $V_G = 0.4V$, source to drain is from left to right, c.f. Figures 5.19 and 5.20. Current density in $A\mu m^{-2}$.

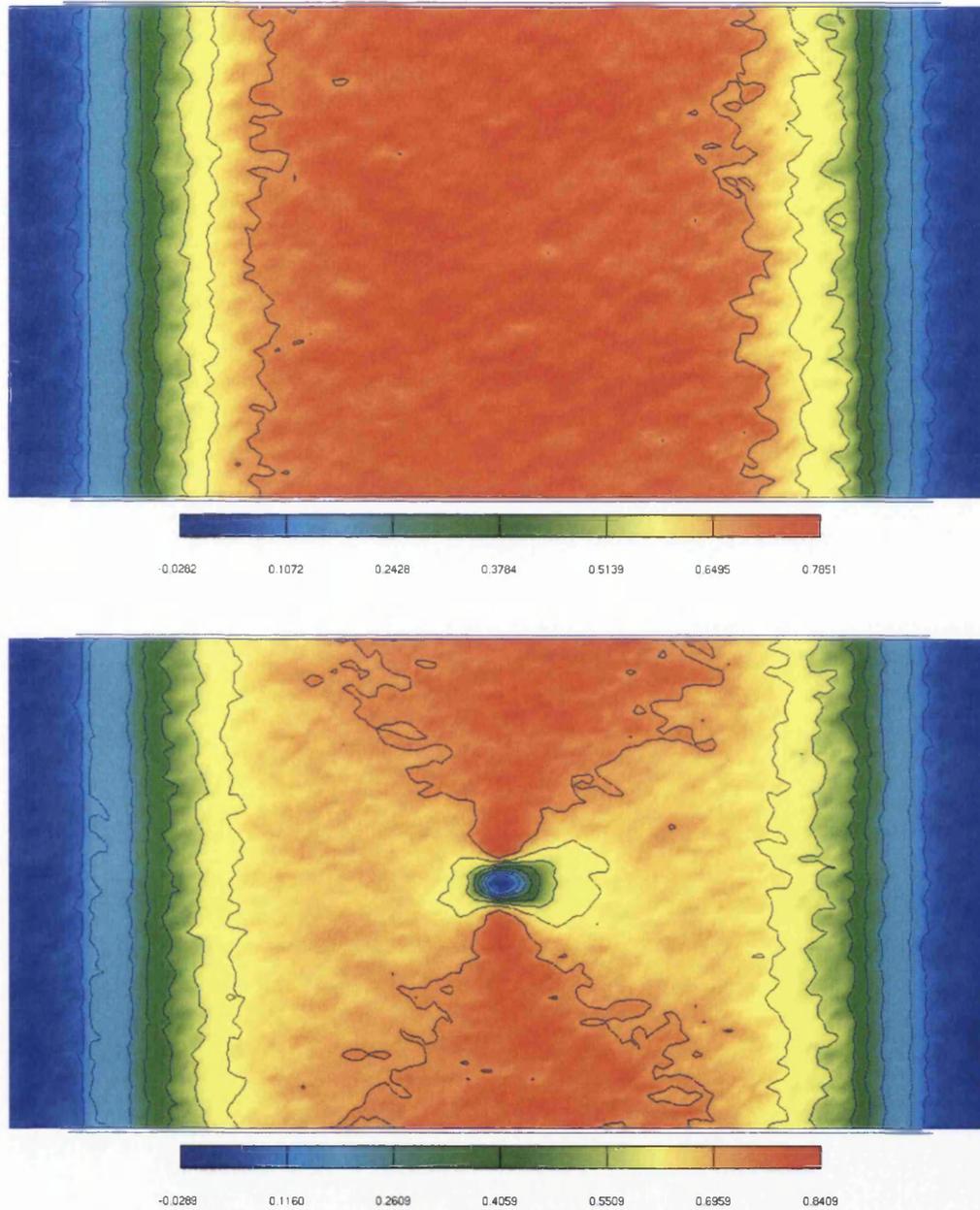


Figure 5.19: x -component of the average electron current density at the interface both omitting (top) and including (bottom) the trapped charge. $V_G = 1.0V$, source to drain is from left to right, c.f. Figures 5.18 and 5.20. Current density in $A\mu m^{-2}$.

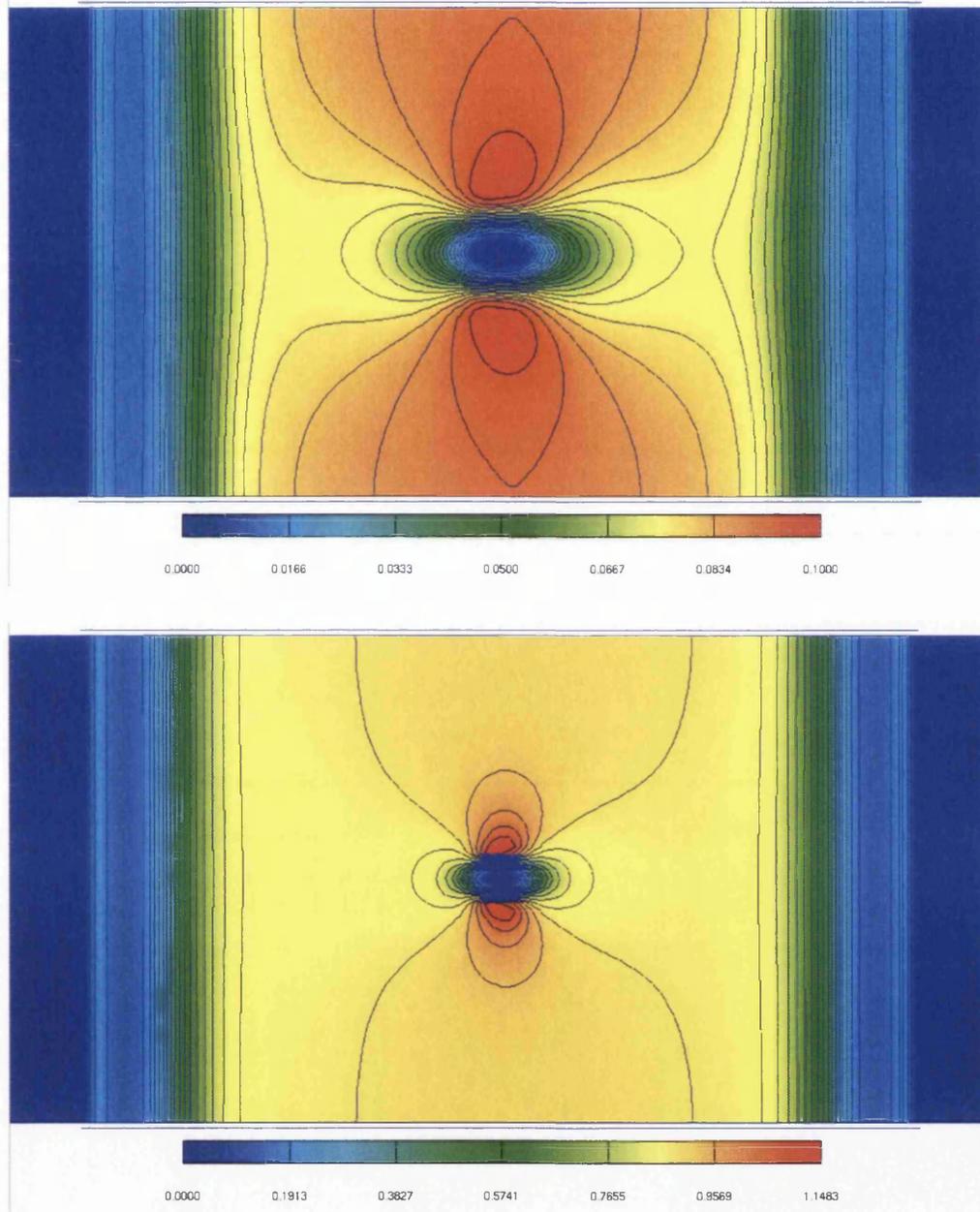


Figure 5.20: x -component of the average electron current density at the interface from DD simulation including the trapped charge. $V_G = 0.4V$ (top), $V_G = 1.0V$ (bottom), source to drain is from left to right, c.f. Figures 5.18 and 5.19. Current density in $A \mu m^{-2}$.

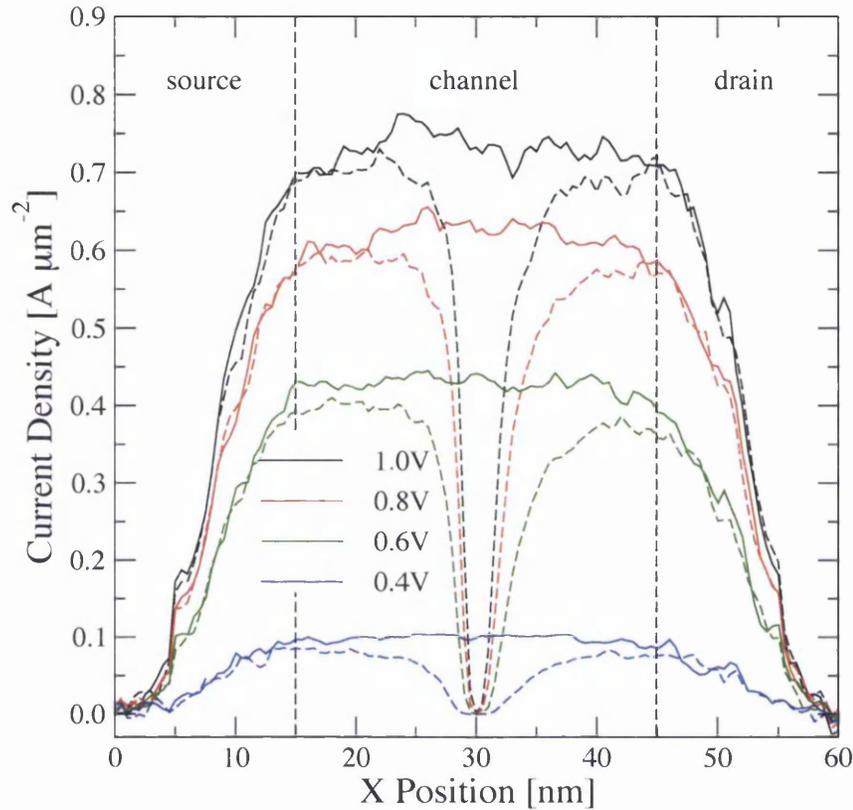


Figure 5.21: Current density profiles along a line perpendicular from source to drain through the position of the trapped charge, both omitting (solid lines) and including (dashed lines) a trapped charge. c.f. figures 5.18 and 5.19. Results plotted for $V_G = 0.4, 0.6, 0.8, \& 1.0V$.

current density is clearly seen and drops to zero at the position of the trapped charge in all cases. What can also be made out now is a lack of symmetry about this point; the current density drops sharply on the source side while it rises more slowly on the drain side. This is due to a lack of symmetry in the electron concentration and is a result not reproduced within DD. The electron density profile from DD and MC simulations along the same line for $V_G = 0.4V$ (where the screening length is greatest) is plotted in figure 5.22. It can be seen that the electron concentration from DD is symmetric about the trapped charge, while that from MC simulation reveals a build up of electrons on the source side and

a depletion on the drain side. It should be noted that the electron concentration from MC simulation is lower than that of DD within the channel due to the differing charge assignment schemes employed. The CIC assignment of the electrons in MC spreads the charge and within the channel does not accurately represent the exponential variation in electron concentration inherent to the DD solution.

This effect on the electron concentration is due to the Coulomb scattering. The repulsive nature of the interaction accelerates electrons away from the trapped electron, but the lateral field from source to drain ensures that more electrons are driven towards the trapped charge from the source side than from the drain side. The deceleration of electrons on the source side result in the accumulation, while those electrons scattered past the trapped charge are additionally accelerated towards the drain.

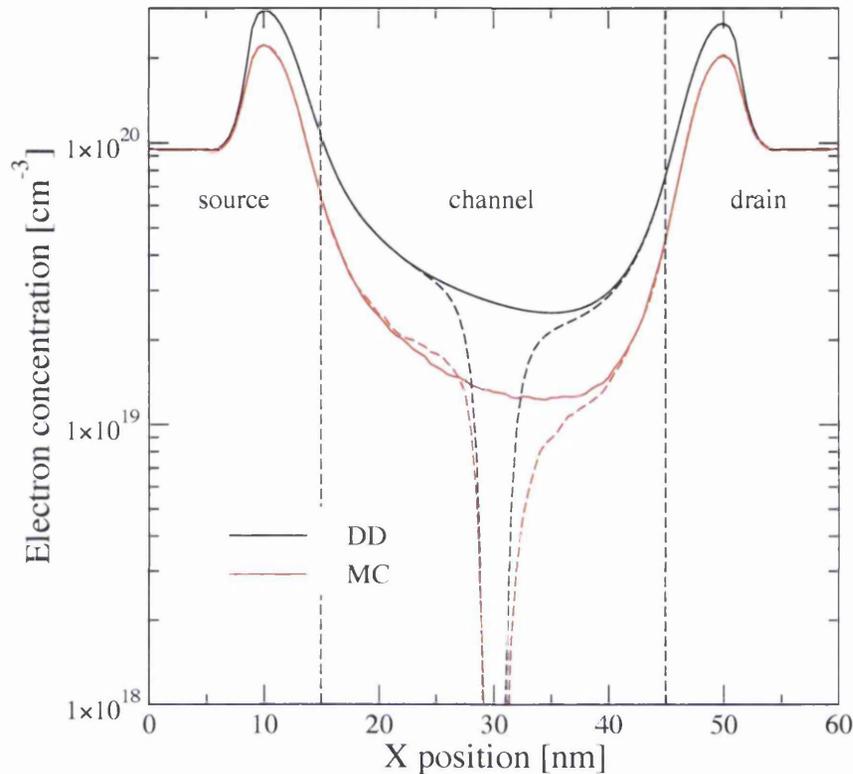


Figure 5.22: Electron density profiles along a line perpendicular from source to drain through the position of the trapped charge, both omitting (solid lines) and including (dashed lines) a trapped charge. $V_G = 0.4V$.

5.3.3.3 Percentage Current Reduction in 20nm and 10nm MOSFETs

The simulation results for the 20nm and 10nm MOSFETs are presented in figures 5.23 and 5.24 respectively and show qualitatively similar trends to that of the 30nm device (figure 5.17). In successively smaller devices, the percentage reduction in current is progressively higher while the contribution from scattering to the total reduction also increases. The same dependence of the electrostatic and scattering contributions as a function of the applied gate voltage is still maintained, with scattering accounting for the majority of the current reduction at high gate voltages while the electrostatics dominate at low gate voltages. It is worth noting that in the case of the 10nm MOSFET the overall drive current

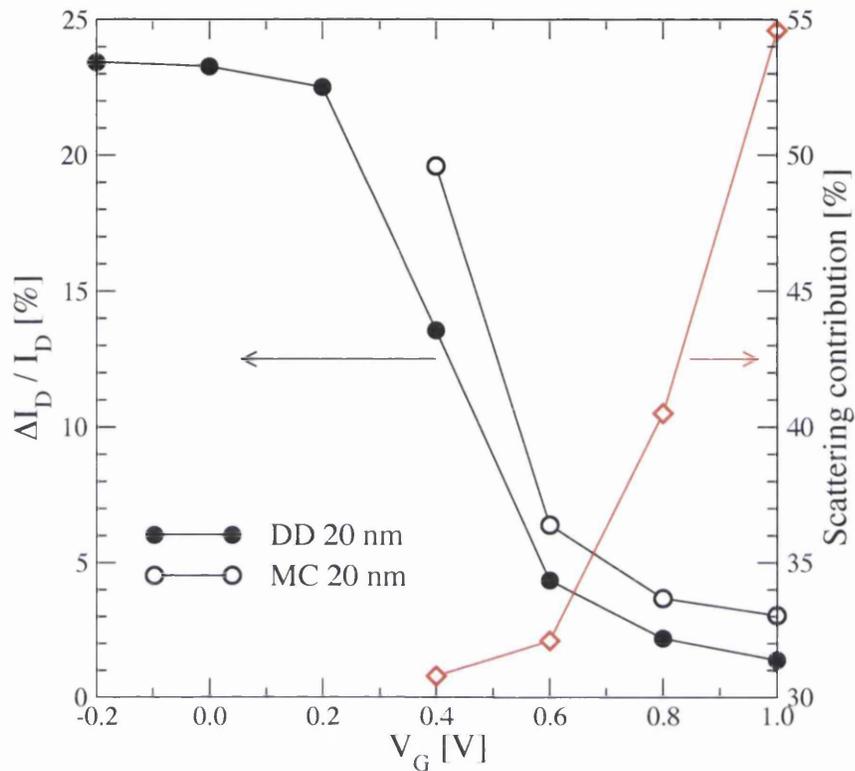


Figure 5.23: Percentage reduction in current upon introducing a trapped charge in the centre of the 20nm MOSFET for a series of applied gate voltages. Results are shown for both MC and DD simulation, showing increased reduction from MC simulation and over the 30nm device 5.17.

reduction associated with the trapping of a single electron is between 8% to 14% at the expected supply voltages of 0.6V to 0.8V for the corresponding technology node.

5.3.4 Conclusion

RTS amplitudes resulting from the trapping of a single electron in the centre of the channel of nano-scale *n*-MOSFETs were simulated using both self-consistent DD and frozen-field Monte Carlo simulations. In the case of continuous doping without considering a trapped charge, the simulation approaches give comparable results when a low drain bias is applied. A series of scaled devices with channel length of 30nm, 20nm and 10nm were

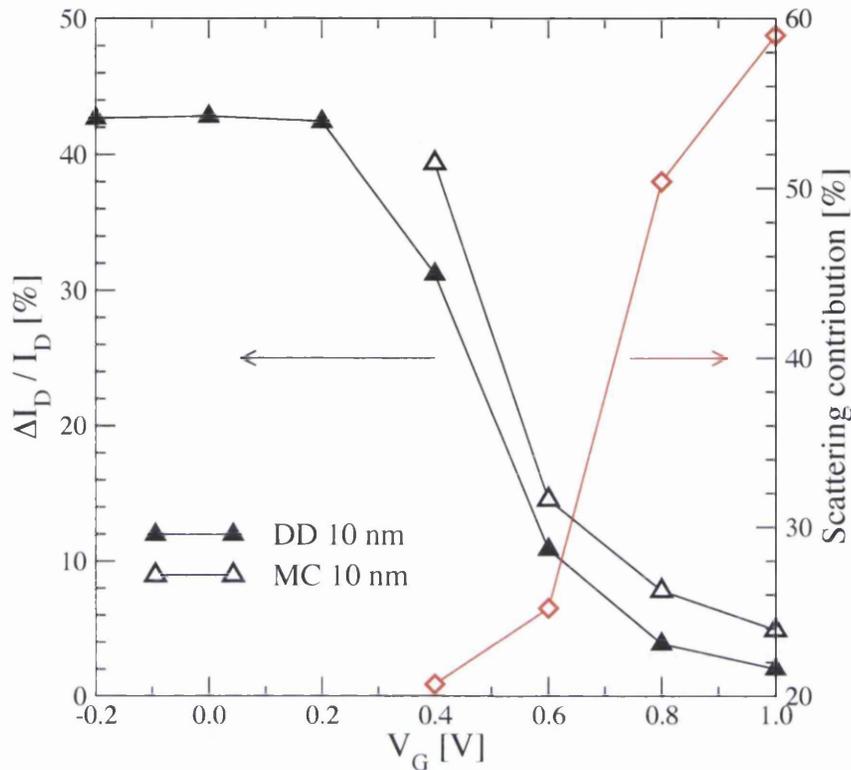


Figure 5.24: Percentage reduction in current upon introducing a trapped charge in the centre of the 10nm MOSFET for a series of applied gate voltages. Results are shown for both MC and DD simulation. Larger reductions in current are seen as compared with figures 5.17 and 5.23 for the 30nm and 20nm devices respectively.

studied over a range of applied gate voltages at a fixed drain bias of 50mV. Compared with DD simulation, MC results in a larger reduction in current as a result of additional Coulomb scattering associated with the trapped charge, an effect unaccounted for in DD. This highlights the importance of correctly incorporating scattering variation, and consequently mobility variation, in nano device simulation.

The contribution of scattering to the total reduction in current from MC simulations increases from around 20 – 30% near threshold, to 40 – 60% in strong inversion. This indicates that the electrostatic influence of the trapped charge is primarily responsible for the reduction in current at low gate bias, while additional scattering becomes equally

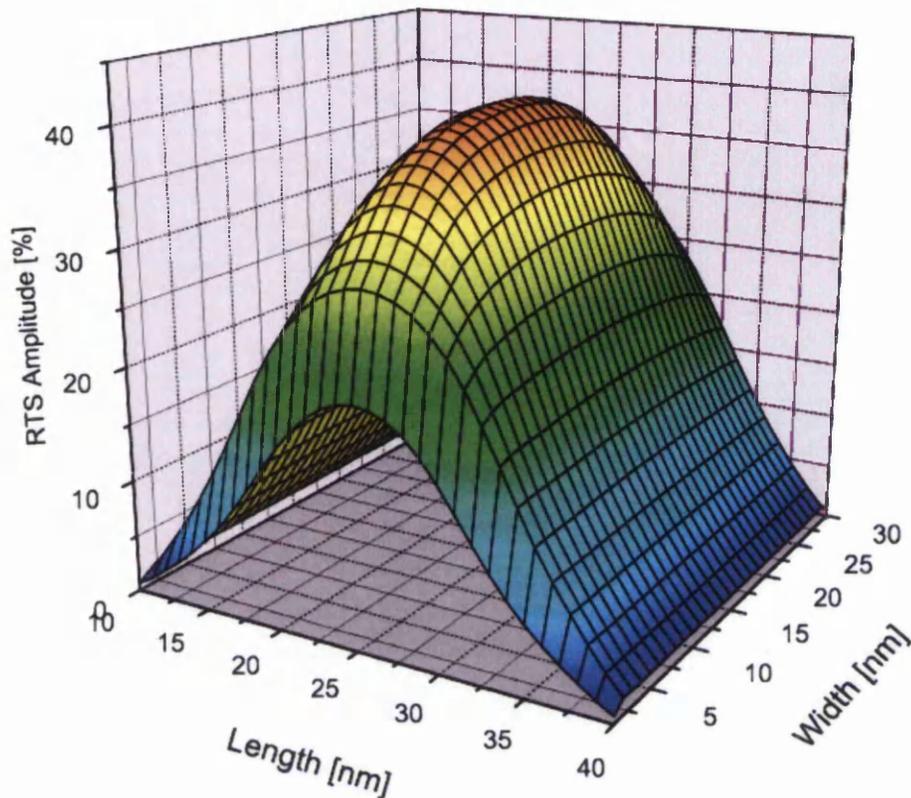


Figure 5.25: Taken from [152], this shows the reduction in current due to a trapped charge placed at varying positions along the semiconductor/insulator interface. The results, obtained from DD simulation, show only the electrostatic influence which is symmetric about the maximum impact in the very centre of the channel.

important or more so at high gate bias. The simulations indicate that the overall reduction in drive current as a result of only one electron trapping can reach more than 10% in a $10\text{ nm} \times 10\text{ nm}$ n -MOSFET.

In all simulations performed, the trapped charge was located in the centre of the channel where, at low drain voltage, it has the greatest influence [152]. Electrostatically in DD simulations this is certainly true, as shown in figure 5.25, because at low drain voltage the maximum of the potential barrier for electrons between the source and drain, which determines the magnitude of the drain current, is in the middle of the channel. A charge trapped

in the centre has the most profound effect on the barrier and therefore the strongest influence on the current. The impact of the trapped charge is also least screened from electrons within the source and drain. Results shown within this section then represent the electrostatic impact at it's peak. However, Coulomb scattering need not necessarily follow the same trend, particularly at high drain voltage. Coulomb scattering is far more effective at low energies, and MC simulation has shown that a scattering centre placed nearer the source/channel junction has a greater effect on current reduction than one placed nearer the drain [153]. This is due to the lower energy of the carriers injected into the channel compared to those accelerated further down the channel. Therefore the relative contributions of the electrostatic and transport effects reported above is not the same for different positions of the trapped charge and at different bias conditions. It should be expected that Coulomb scattering will be more of a dominating factor when the trapped charge is located near the source. Here scattering is most important while the electrostatic influence could be reduced due to the screening from the heavily doped source region. Moving towards the drain, particularly at high drain voltage, it is unclear as to whether or not Coulomb scattering will remain the dominant factor reducing the current as the ability to scatter is also reduced.

5.4 Atomistic MOSFET

In the previous section, the importance of accounting for variation in transport due to Coulomb scattering from an individual discrete charge was highlighted. It was shown that MC simulations result in a significant increase in current variation in the presence of a discrete charge compared to DD simulation results. While in the previous simulations the scattering from the trapped charge was treated directly, the ionized impurity scattering from the otherwise uniformly doped channel was treated through the Brooks-Herring

scattering rate. Naturally, in such small devices, the ionized impurity scattering resulting from the discreteness of the doping charge should be similarly treated in an *ab initio* manner.

The statistical analysis of device parameter variations has previously been performed using the DD approach, which is still the most computationally efficient method to study a large ensemble of devices. Therefore it will be very interesting to compare intrinsic parameter variations for a fully atomistic device for an ensemble of devices obtained from DD simulations and MC simulations incorporating *ab initio* Coulomb scattering. It is expected that the proper accounting of the position dependent Coulomb scattering, and hence mobility variation, compared with the simple mobility models employed within the DD simulation will lead to larger device parameter variations. The magnitude of this difference and consequently the limitations of the DD simulation used for statistical analysis of random dopant induced parameter variations is the subject of the next section

5.4.1 Introduction

Previous simulation studies of random dopant fluctuation effects, similar to the studies of effects associated with trapped charges, have been predominantly performed using 3D Drift-Diffusion (DD) simulators [46, 54, 55, 65]. Again, such simulations only capture the electrostatic effects associated with random discrete dopant distributions in providing an estimate for the variations in the threshold voltage and the drive current. However, the DD simulations cannot capture the complex effects associated with the variation in carrier transport, from device to device, associated with the different numbers and configuration of ionized dopants within the channel that act as Coulomb scattering centres in a non self-averaging manner. Therefore the published results for the variation in the drive current obtained from DD simulation most probably underestimate the real magnitude of the

intrinsic parameter variations.

Although *ab initio* Coulomb scattering from ionized impurities through the real space carrier trajectories has previously been included in 3D ensemble Monte Carlo simulations [77, 79, 139, 142], little statistical analysis has been performed with such simulators. In particular, no results have been published comparing the impact of the variation of numbers and position of Coulomb scattering centres in the channel on the current fluctuations, with their electrostatic impact on the current percolation paths within the device. In this section, using a methodology similar to the methodology used in the previous section, a careful comparison between DD and MC simulations of MOSFETs with different discrete dopant configurations is presented. The relative importance of the electrostatic and scattering effects when studying random dopant induced intrinsic parameter variations is investigated and highlighted.

5.4.2 Simulation Methodology

The variation in current associated with a random distribution of dopant atoms within the active region of the $30\text{nm} \times 30\text{nm}$ *n*-MOSFET simulated previously was studied in comparison with simulation of a uniformly doped device of the same design. An ensemble of 50 microscopically different but macroscopically identical devices were randomly generated and simulated using the DD simulator. These devices have continuously doped source and drain regions while the channel region is described atomistically. In the DD simulations, as before, field dependent and concentration dependent mobility models are employed, though in the atomistic DD simulations the mobility values are taken from values stored at each mesh node as a result of simulation of the continuous device. In all simulations, the gate bias is set at 1.0V while the source to drain bias is 50mV , low in order to minimize non-equilibrium transport effects and allow fair comparison between

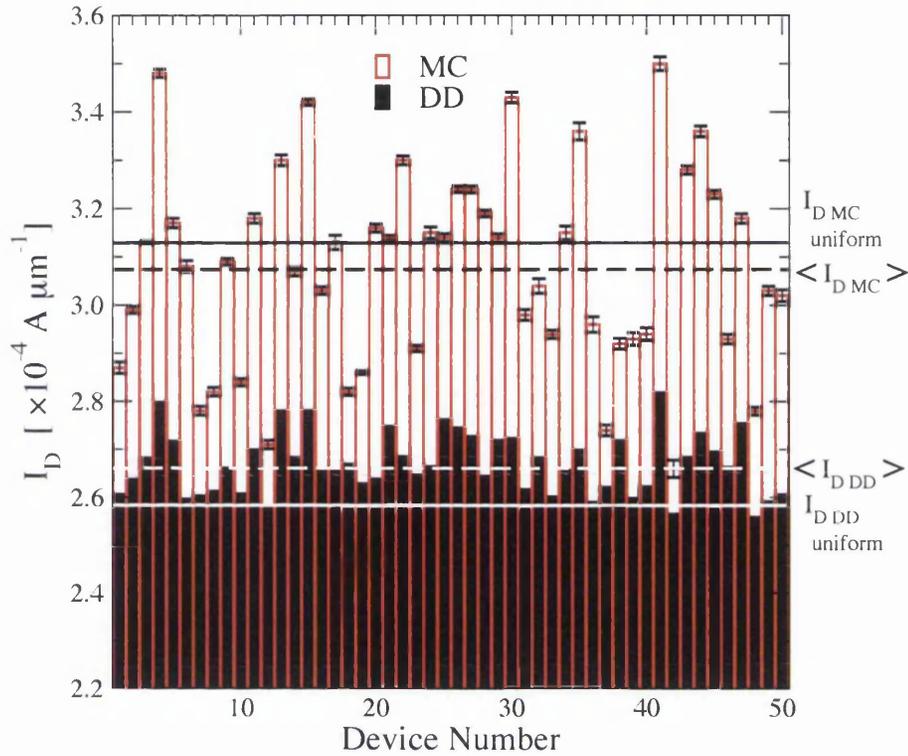


Figure 5.26: Histogram of the drain-current for each of the 50 randomly configured devices as calculated from DD and MC simulation. MC yields larger currents but reproduces the trend of DD simulation.

the two simulation methods. MC simulation is then performed in the frozen-field approximation. The MC and DD simulation of the atomistic devices, in comparison with their respective results for the continuous device, allow the variation in current to be determined for each atomistic device and comparison between DD and MC results to be made.

5.4.3 Results

5.4.3.1 Current distribution

The drain currents for the 50 randomly generated devices obtained from both drift diffusion and Monte Carlo simulation are shown in figure 5.26. The error in the esti-

mation of the current from Monte Carlo is plotted along with the currents and is typically less than $\pm 0.5\%$. For all devices, the Monte Carlo simulations result in larger currents when compared with the equivalent drift diffusion result, and the mean current over the 50 devices is an additional 15% over the mean from drift diffusion. Respectively, the values for the mean drain currents from drift diffusion and Monte Carlo are $\langle I_{D DD} \rangle = 2.67 \times 10^{-4} A \mu m^{-1}$ and $\langle I_{D MC} \rangle = 3.08 \times 10^{-4} A \mu m^{-1}$ and are plotted as the white and black dashed lines as marked in figure 5.26. As well as plotting the mean atomistic current, the drain current from DD and MC simulations of the uniformly doped device are also plotted as the solid black and white lines, as marked, and are equal to $I_{D DD uniform} = 2.59 \times 10^{-4} A \mu m^{-1}$ and $I_{D MC uniform} = 3.13 \times 10^{-4} A \mu m^{-1}$ for DD and MC respectively. From this it can be seen that the atomistic DD simulations result in an average current greater than the uniform case, while MC simulation shows a reduction in the average current. Despite this difference, the trend in the increase and decrease of the current between devices is largely reproduced between both simulation models. This is clearly seen in the plot of I_D from DD Vs. I_D from MC for all 50 devices, which is shown in figure 5.27. The currents are correlated, with a calculated correlation coefficient of 0.80, while the spread highlights the transport variations imposed by the random dopants and captured within MC simulation over and above the electrostatic effect responsible for the current variation in DD.

5.4.3.2 Percentage Change in Current

Figure 5.28 shows the distribution in the percentage change in current from the uniform device for both DD and MC. From the DD results it can be seen that, as previously seen in figure 5.26 and agreeing with the well known result [47], the atomistic description increases the average current due to the presence of current percolation paths. This is a purely electrostatic effect. However, this trend is again clearly not seen to be reproduced

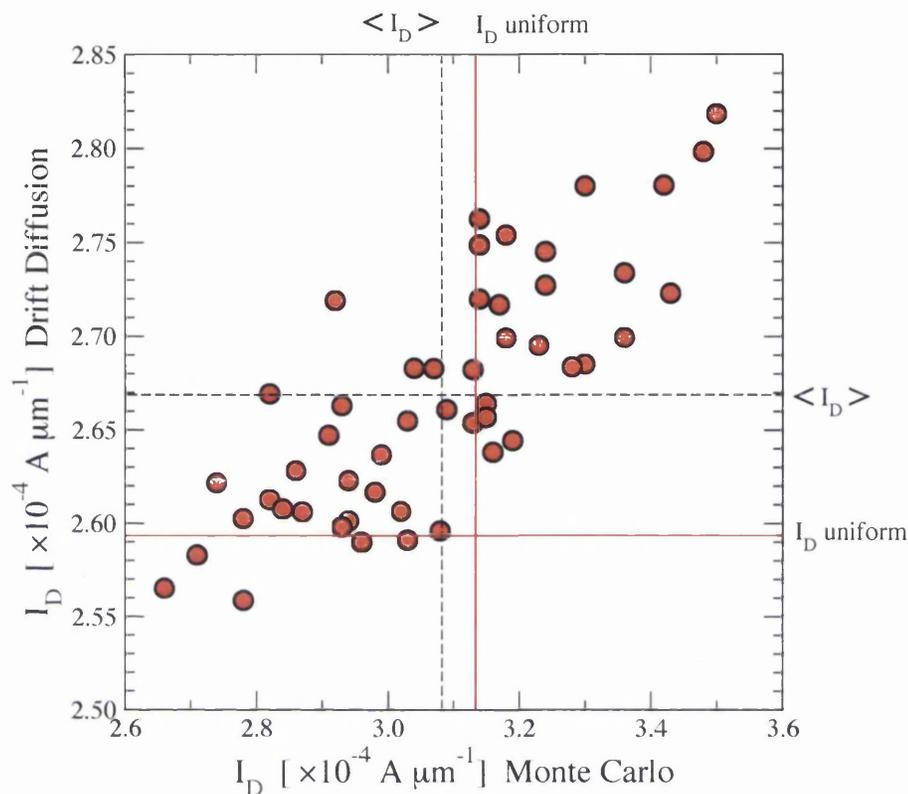


Figure 5.27: Scatter plot of drain current calculated from DD against that calculated from MC. The correlation coefficient for this distribution is calculated as 0.80. Shown again is the average current and current from uniform simulation for both simulation techniques.

in the MC simulations. 26 of the 50 devices show reduced current in the MC simulations while only 5 show reductions in DD. And while the reductions in current from DD are relatively small, the reductions in MC are seen to be comparable with the largest increase in current seen in MC. The percentage variation in current from DD is plotted against that from MC in figure 5.29. Again, the correlation between the results is apparent.

At this point it may be necessary to question the size of the statistical sample. The difference between the current from uniform DD simulation and the average atomistic DD current is $\approx +3.0\%$ while the same difference from MC simulation is $\approx -1.5\%$. This is a small difference, but the distribution of currents from DD clearly favour an increase

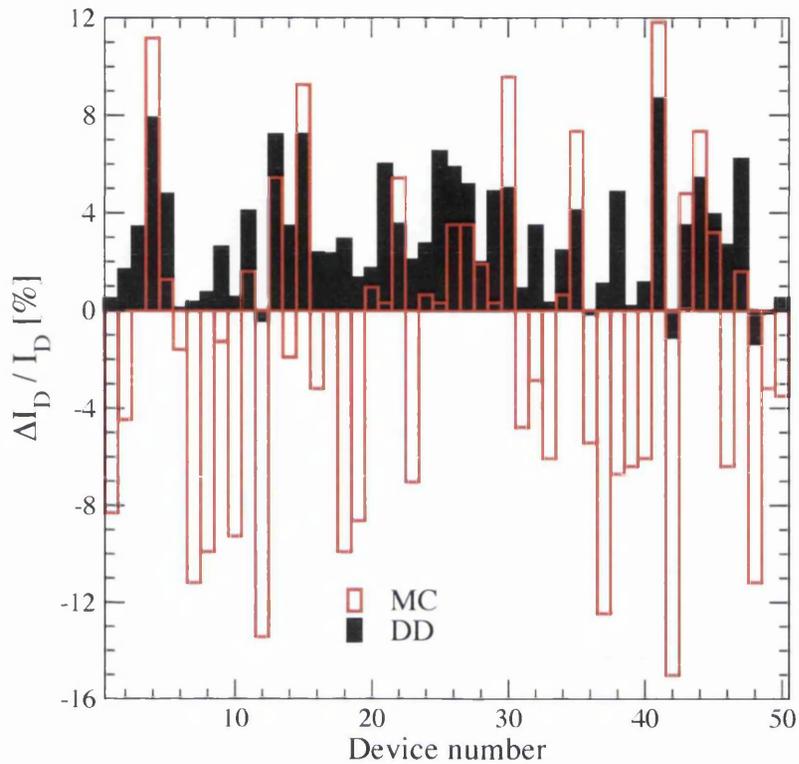


Figure 5.28: Histogram of the percentage change in current for each of the 50 randomly configured devices. MC simulations show a large number of devices with a significant reduction in current, contrary to DD.

while those from MC are evenly distributed about the uniform device. From the limited set of data it is inappropriate to claim that MC simulation results in an average reduction in current since just a few devices with increased current may change this. What is clear however, is that the magnitude of the current variation is larger than compared to DD. The standard deviation in the current calculated from DD is $6.4 \times 10^{-6} A \mu m^{-1}$ which is 2.4%, while the standard deviation from MC is $2.1 \times 10^{-5} A \mu m^{-1}$ which is 6.7%.

5.4.3.3 Drift Diffusion Vs. Monte Carlo

From figure 5.27 we take 3 devices with the same current from DD simulation but with a spread in values for the current estimated by MC. These are devices 29, 30 and 38.

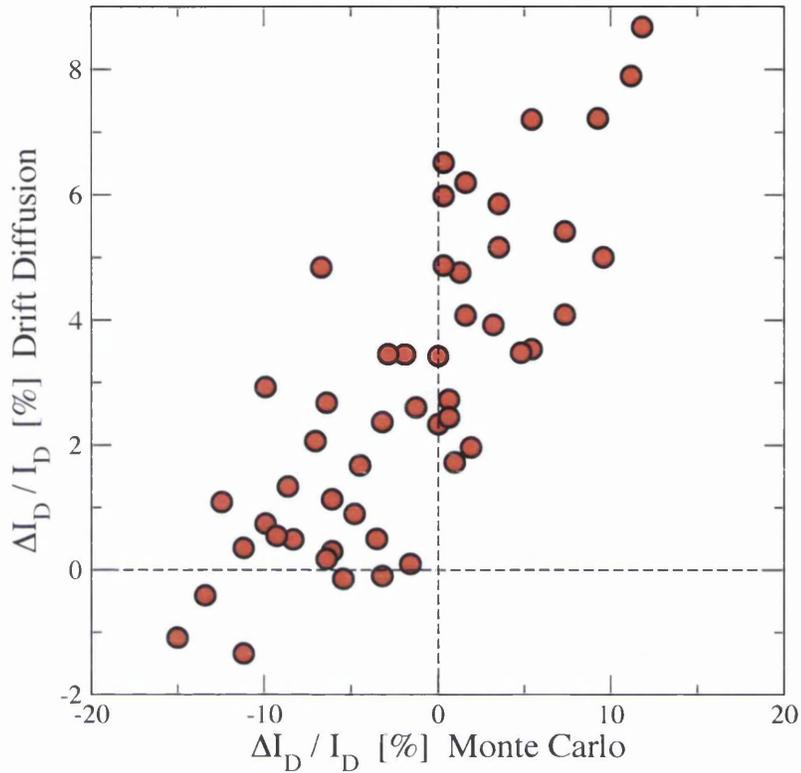


Figure 5.29: Scatter plot of the percentage change in current calculated from DD against that calculated from MC. MC simulations show similar spread of values about 0, contrary to DD.

The currents for each of these devices is tabulated in table 5.4. In order to explain the differences in current, the current density at the interface of each device from DD is compared to that of MC and to each other.

Device 30: Figures 5.30 and 5.31 show the component of the current density from source to drain (left to right) at the interface from DD simulation and the corresponding frozen field MC simulation respectively. Good qualitative agreement is seen between the two distributions, with the regions of lowest current density localised about the individual acceptor atoms. There are 3 acceptors near the interface in the channel at the top of the figure, and a further 2 acceptors buried near the interface in the middle of the

Device		30	29	38
$I_D (DD)$	$\times 10^{-4} A \mu m^{-1}$	2.723	2.719	2.719
$\Delta I_D / I_D (DD)$ [%]		5.01	4.88	4.88
$I_D (MC)$	$\times 10^{-4} A \mu m^{-1}$	3.43	3.14	2.92
$\Delta I_D / I_D (MC)$ [%]		9.59	1.92	-6.71

Table 5.4: Devices 29, 30 and 38 have similar values for I_D from DD simulation, but those from MC simulation show large variation.

figure that clearly contribute to affect the current. The relatively low number of dopants in this device, and their close grouping, creates a large un-doped area in which the current density remains comparatively high in the MC simulation. Within DD the current density similarly peaks around the acceptors but rapidly falls off and does not show the same high level in the un-doped regions. This results in the large percentage increase in the current seen in MC simulation.

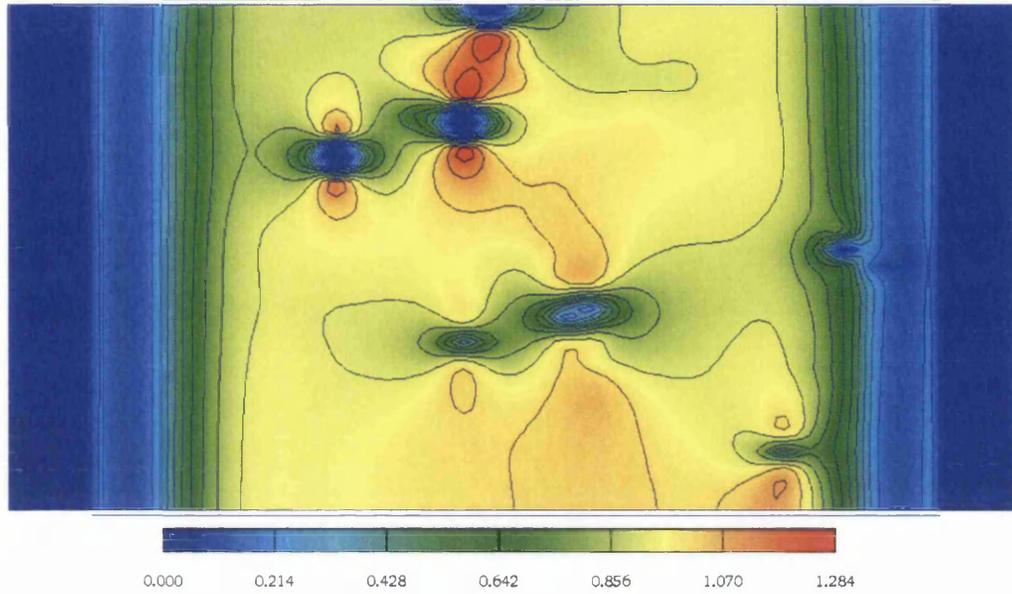


Figure 5.30: Component of current density along source to drain (left to right) at the Si/SiO₂ interface from drift diffusion simulation of device 30. Current density in $A \mu m^{-2}$.

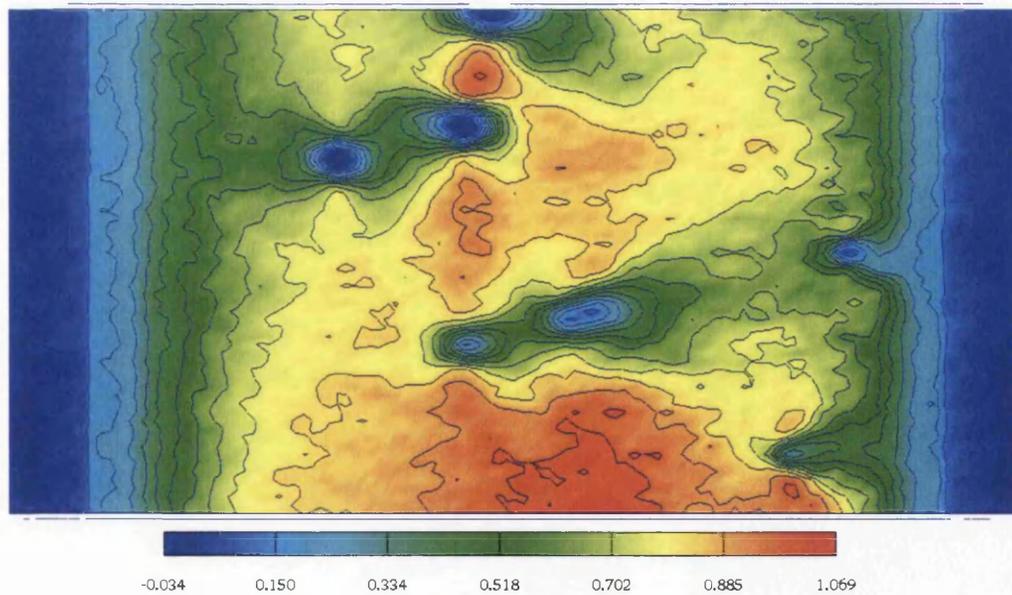


Figure 5.31: Component of current density along source to drain (left to right) at the Si/SiO₂ interface from Monte Carlo simulation of device 30. Current density in $A \mu m^{-2}$.

Device 29: Comparing MC simulations, device 30 has lower current than device 29. The current density component from source to drain resulting from DD and MC is shown for device 29 in figures 5.32 and 5.33 respectively. Within this device there are far more acceptor atoms evident within the channel, though the majority are not at the interface and hence their effects are somewhat lessened. Within DD, the increased number of acceptors is compensated for by their position deeper in the channel resulting in little change in the current, while within MC simulation the acceptors are distributed so to remove the kind of large un-doped regions that resulted in the large current seen in device 30. This is responsible for the reduced current found in MC. Again, the current in the regions between acceptors remains relatively large in the MC simulation while it falls off in DD.

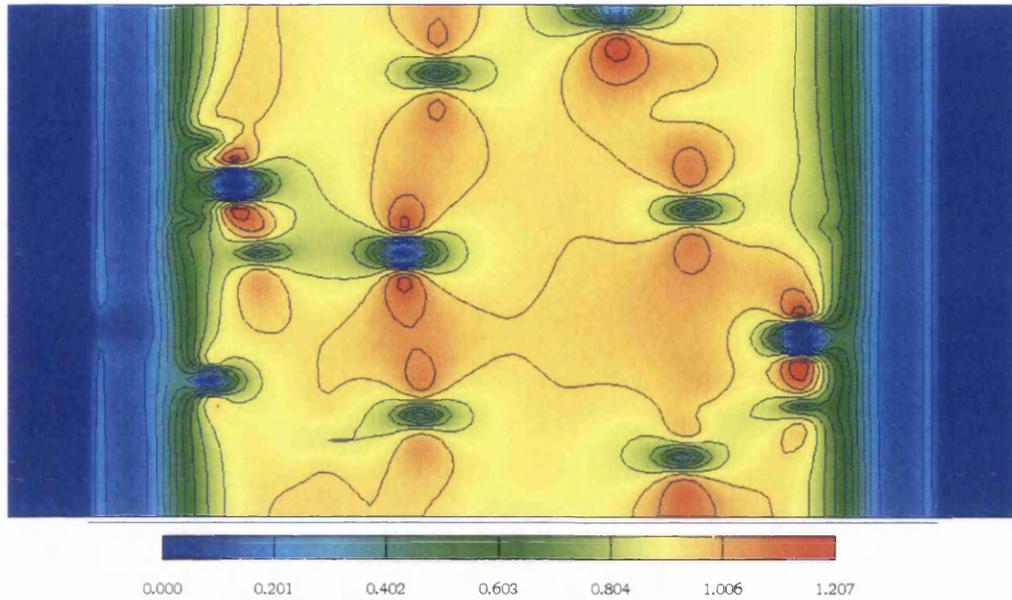


Figure 5.32: Component of current density along source to drain (left to right) at the Si/SiO₂ interface from Drift Diffusion simulation of device 29. Current density in $A\mu m^{-2}$.

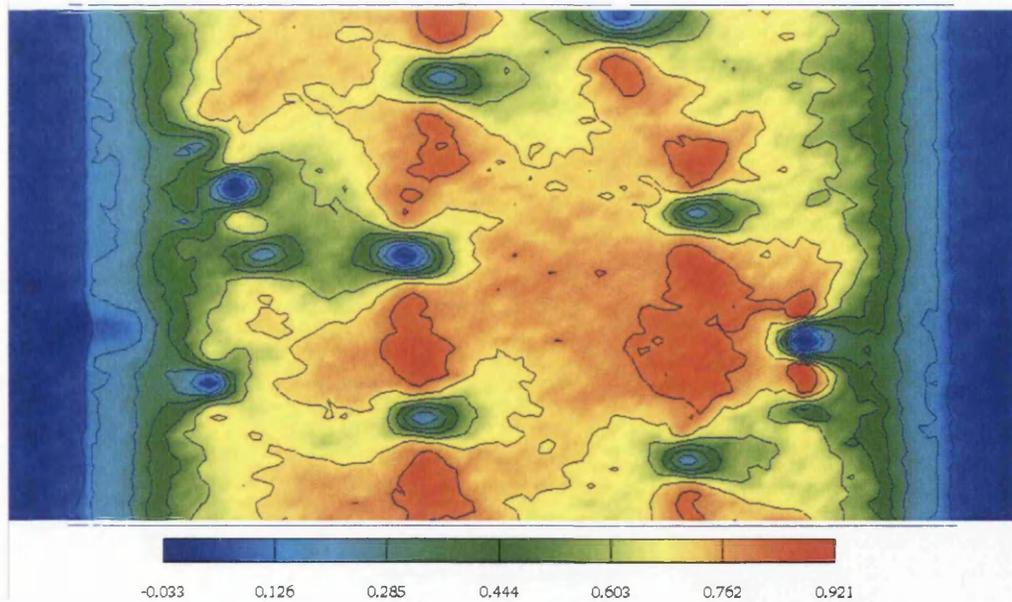


Figure 5.33: Component of current density along source to drain (left to right) at the Si/SiO₂ interface from Monte Carlo simulation of device 29. Current density in $A\mu m^{-2}$.

Device 38: Device 38 has the lowest current of the three from MC. Figures 5.34 and 5.35 show the current density from DD and MC for this device. Compared to device 29 there are fewer acceptors influencing the current within the channel, but there are four present at the interface and distributed across the width of the channel. Two acceptors are also seen at the interface in the source region. In this case the positioning of the dopants, rather than the number, primarily effect the current in MC. The lines of dopants across the width of the channel efficiently impede current flow and are responsible for the reduced current. Comparing with device 30, there are similar numbers of acceptors at the interface which act to reduce the inversion carrier charge density, hence similar currents are observed from DD

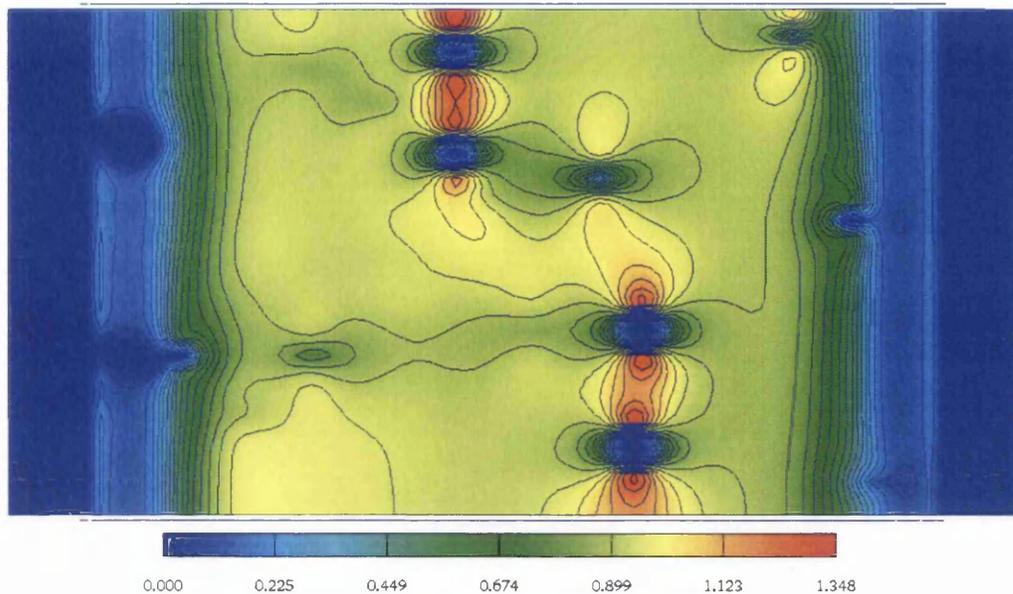


Figure 5.34: Component of current density along source to drain (left to right) at the Si/SiO₂ interface from Drift Diffusion simulation of device 38. Current density in $A\mu m^{-2}$.

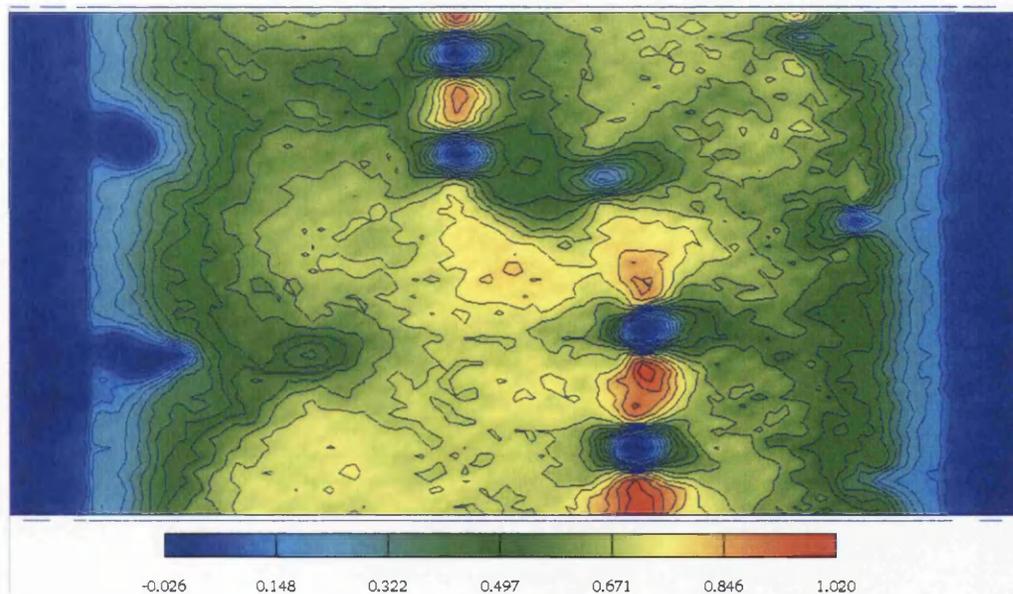


Figure 5.35: Component of current density along source to drain (left to right) at the Si/SiO₂ interface from Monte Carlo simulation of device 38. Current density in $A\mu m^{-2}$.

5.4.3.4 Self-consistent Monte Carlo simulation

The largest current in both DD and MC is associated with device 41. In order to estimate the validity of the frozen field approximation in the atomistic devices, a self-consistent Monte Carlo simulation of device 41 was performed. The simulation was performed under the same conditions as the frozen field simulations. A total of 80,000 particles were used to sample the initial electron concentration from DD and but were only simulated for 100,000 time steps of length $1 \times 10^{-16}s$.

From frozen field simulation the drain current was estimated as

$$I_D = (3.50 \pm 0.02) \times 10^{-4} A \mu m^{-1}$$

Self consistent simulation gave an estimated current of

$$I_D = (5.51 \pm 0.02) \times 10^{-4} A \mu m^{-1}$$

a 57% increase over the frozen field estimation. The current density at the interface from both frozen field and self-consistent MC simulations are shown in figures 5.36 and 5.37 respectively, and show a qualitative agreement. The self consistent simulation shows higher current density, consistent with the increased drain current, and is found to be linked with increased electron density within the channel. This is as a direct result of the self consistency better maintaining the electron concentration within the extended source region and subsequent injection into the channel. The frozen field accurately maintains the concentration within the contact at the point of injection itself, but through diffusion alone can not as effectively maintain the concentration injected into the channel. This may likely be a systematic error in the frozen field simulations since the source regions are identical. But while the currents from frozen field are likely to be underestimated compared to self

consistent simulation, the variation in current may be consistently reproduced. However, that this is the case would require self consistent simulation of the entire ensemble of devices for comparison.

As well as comparing the current density, comparison of the potential from DD, shown in figure 5.38, and the time averaged self consistent potential from MC, figure 5.39, again shows qualitative similarities. This further lends credence to the validity of the frozen field MC simulations

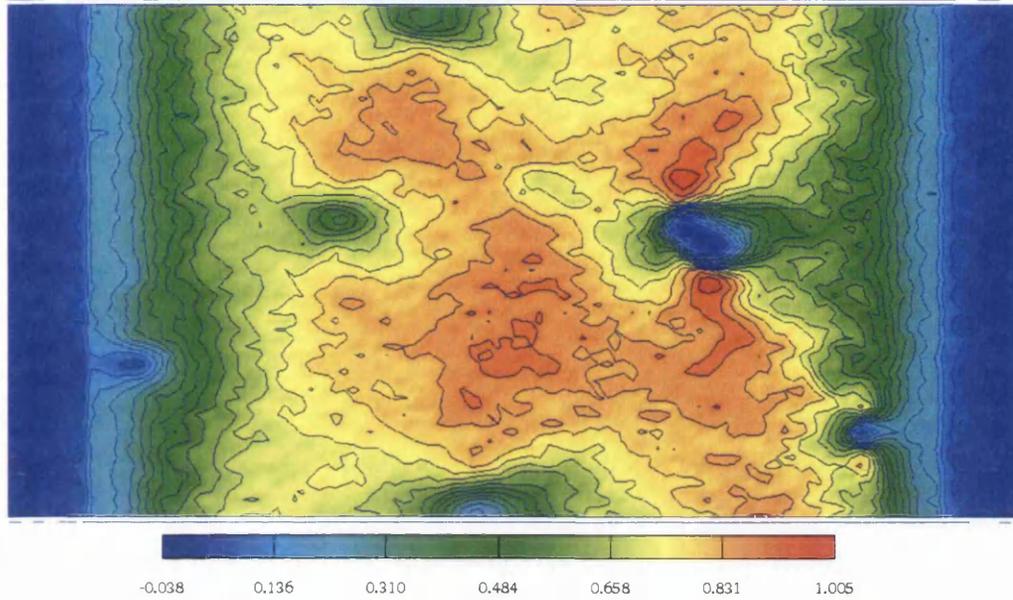


Figure 5.36: Component of current density in $A \mu m^{-2}$ along source to drain (left to right) at the Si/SiO₂ interface from frozen field Monte Carlo simulation of device 41.

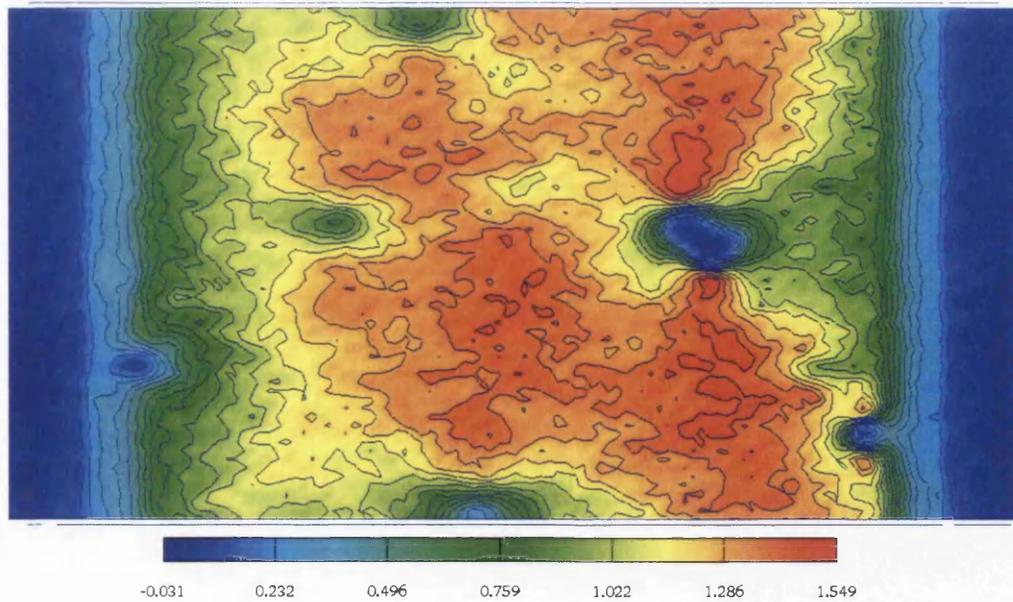


Figure 5.37: Component of current density in $A \mu m^{-2}$ along source to drain (left to right) at the Si/SiO₂ interface from self consistent Monte Carlo simulation of device 41.

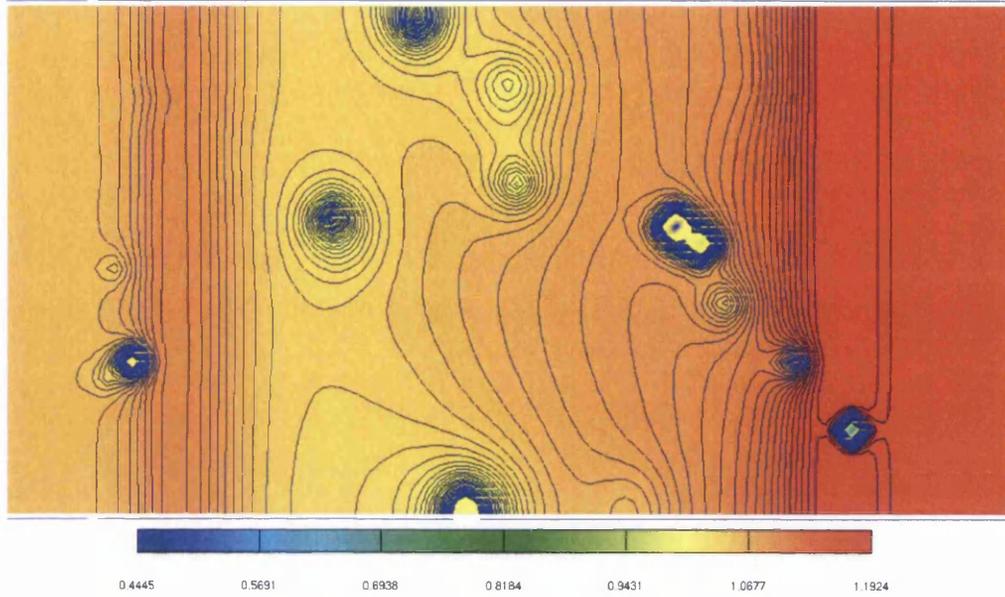


Figure 5.38: Potential profile from DD solution of device 41 taken at the interface. Potential in Volts

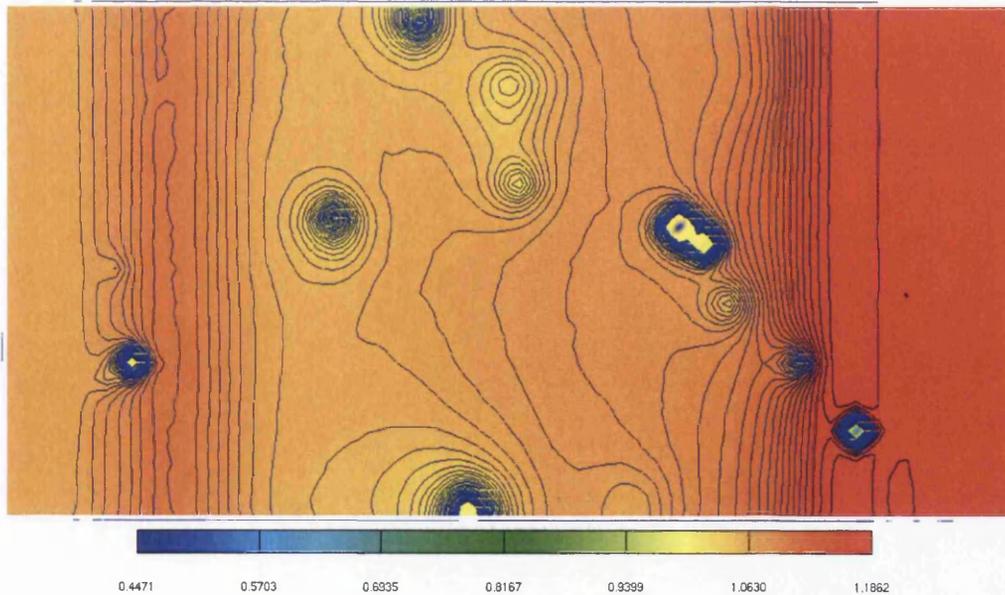


Figure 5.39: Potential profile from self consistent MC solution of device 41 taken at the interface. potential in Volts, c.f. Figure 5.38 above

5.4.4 Conclusion

The inclusion of position dependent ionized impurity scattering within MC simulation of an ensemble of 30nm n -MOSFETs shows increased current variation compared with the simulation of the same ensemble performed using DD. The DD results show an average increase in current of 3.0% with a standard deviation of 2.4% . Such simulations only consider the electrostatic influence associated with unique charge configurations. By comparison, MC simulation results in an average current decrease of 1.5% with an increased standard deviation of 6.7% . The trend in current variation between devices within the ensemble from DD are reproduced within the MC results, but the position dependent impurity scattering from the unique arrangements of dopants reduces the correlation. The comparison of the current variation obtained by both methods gives a correlation coefficient of 0.80 . Although the electrostatic impact of a particular configuration of donors may be similar, differences in the position dependent scattering between such configurations may have larger effects. By effecting both the mobility and the electron concentration throughout the channel, the ionized impurity scattering plays an important role in determining the magnitude of the random dopant induced device parameter variations.

It is noted that the MC simulations reported for the ensemble of devices are performed in the frozen-field approximation to allow for efficient simulation. A low drain bias of 50mV was necessary to best facilitate this approximation. Variation at higher drain bias, particularly in the operating regime, would be a far more important study but requires self-consistent simulation. Under the low bias conditions reported here, it is reasonable to expect that self consistent simulation should also yield differences in the current variation through, in part, the variation in the steady state electron concentration, since the electron concentration is otherwise representative of the initial DD solution which is insensitive to the position dependent scattering. Such self consistent simulations are easily performed

with the developed code, though simulating a statistical set of devices is at present prohibitively time consuming. They would better describe the dynamic screening of dopants by mobile charges under the actions of position dependent scattering, which will in turn alter the scattering potentials.

One such self consistent atomistic simulation was performed for comparison and shows qualitative agreement with the frozen field current density and potential distributions. The simulation however gave a larger current, but on it's own says nothing about the magnitude of the current variation expected. Results presented here highlight the increase in statistical parameter variation from transport variation and the requirement for a proper accounting of ionized impurity scattering in nano scale devices.

5.5 Summary

This chapter addressed the effects of transport variations associated with discrete charges and charge configurations. Within simple 30nm , 20nm and 10nm MOSFET structures the effect of transport variation induced by a single trapped charge was separated from the accompanying electrostatic influence. It was found that additional scattering from the trapped charge increases the current reduction and the corresponding RTS amplitudes over the entire range of applied gate voltages and device sizes. Increased scattering from the trapped charge was a dominant effect at high gate voltage, where electrostatic effects are suppressed, while at low gate voltages the electrostatic charge exclusion around the charge is the dominant current reduction mechanism. The additional scattering from the trapped charge was found to be more important in smaller devices, where the scattering potential necessarily occupies a larger fraction of the channel.

An ensemble of the same 30nm device structure, but with atomistic channel doping, was simulated to compare current variations with and without the inclusion of transport

variation through position dependent Coulomb scattering. It was found that the proper accounting of the Coulomb scattering associated with the unique arrangement of dopants within each device again plays a strong role in determining the current variation by both influencing the mobility and the electron concentration throughout the channel. Larger variations were observed in the MC simulations compared with the DD simulations.

This highlights the importance of properly accounting for the propagation through a random arrangements of scattering centres within nano-scale device simulation. It suggests that previous studies of RTS amplitudes, V_t fluctuations and current fluctuations associated with random dopants, typically performed using DD simulations, underestimate the true magnitude of the statistical variation. A rigorous self-consistent *ab initio* Monte Carlo simulation is a more accurate tool for studying intrinsic parameter fluctuations by properly accounting for the screening from a random arrangement of scattering centres and is required for studies of variation at high drain bias.

Chapter 6

Conclusion

A Monte Carlo simulator was developed that can accurately incorporate Coulomb scattering of carriers from a random configuration of ionized impurities, and the Coulomb interaction mediated carrier-carrier scattering, through the real space trajectories of the carriers themselves. This “*ab initio*” scattering model was thoroughly validated, also carefully checking the carrier distribution and temperature, with respect to bulk mobility simulation. Monte Carlo simulations featuring *ab initio* Coulomb scattering were then used to capture the effect of transport variation on the magnitude of the random dopant induced device parameter variation. In comparison with Drift-Diffusion simulations that only capture the electrostatic variation in charge density, important conclusions about the relative importance of variations in Coulomb scattering were inferred. The proper *ab initio* inclusion of the variation in carrier scattering increases the magnitude of the intrinsic parameter variation and suggests that Monte Carlo simulation is required for a thorough evaluation of such effects.

For clarity, the details of the Monte Carlo simulator are discussed in section 6.1. Evaluation of the *ab initio* scattering method is given in section 6.2 before conclusions related to the application to device variation are presented in section 6.3. The natural extension

of this work is finally discussed in section 6.4.

6.1 The Monte Carlo Simulator

A complete description of the MC simulator was given in chapter 3. An analytical spherical, non-parabolic band model is employed since accurately resolving the high energy electron distribution is unimportant within this work. The considered phonon scattering mechanisms were described and the calibration of the corresponding scattering rates with respect to bulk simulation of field dependent velocity and energy was given. Surface roughness scattering, which is important for realistic MOSFET simulations, is not included in this work, neither is impact ionization which should become important at high fields. The extension of the bulk MC simulator to 3D device simulation was described, detailing the choice of the discretization, charge assignment and force interpolation. All simulations are performed on a uniform rectilinear mesh and charge is assigned via the Cloud-In-Cell scheme. The field at each mesh node is evaluated via a finite difference scheme and the force tri-linearly interpolated to a point within the mesh. The resultant full 3D simulator was verified with the simulation of a simple N-I-N diode using the traditional Brooks-Herring ionized impurity scattering rate.

6.2 *ab initio* Coulomb Scattering

The development of an *ab initio* treatment of Coulomb scattering which replaces the ionized impurity and electron-electron scattering rates in “atomistic” MC simulations was described in Chapter 4. The accurate integration of the electron’s equations of motion was highlighted as paramount to accurate simulation of *ab initio* Coulomb scattering. Consideration was given as to the ability of the numerical integration approaches to con-

serve energy during the close range interaction of the carriers with a point charge and over reflection from a surface, both of which are important within *ab initio* device simulation. The constraints on the numerical integration scheme include (i) only one force evaluation per solution of Poisson's equation to avoid self-forces, (ii) time centred approach in order to approximate a variable mass during propagation in a high field and (iii) synchronisation of the momentum and position components in order to allow for scattering and reflection events. Such requirements were satisfied by the Velocity-Verlet algorithm which has been common to Plasma physics simulation for many years.

However, the above approach to the integration of the equations of motion is still not accurate enough to ensure the accuracy and energy conservation in the infinite potential well of a point charge representing a donor atom within the simulation, particularly at small impact parameters. This prompted a modification of the short-range charge interaction to limit the field and to improve energy conservation. Limiting the field, however, limits the ability of a point charge to act as a scattering centre and could affect the accuracy of the *ab initio* approach. A variety of short-range interaction models were investigated, comparing both the energy conservation and the proper reproduction of the classical scattering angle dependence upon impact parameter. Study of the impact of the short-range approximations has not been performed elsewhere. It was found that an analytical short range interaction model in which the Coulomb force reaches a maximum value at a chosen cut-off radius before dropping to zero at zero separation, delivers a vast improvement in energy conservation compared to the simpler models adopted in literature while providing a similar scattering angle dependence. This novel model was then adopted for the rest of the work.

The application of the short-range force correction in full scale 3D simulations was further described. The corresponding *ab initio* approach was verified through the reproduction of the concentration dependence of the low field electron bulk mobility in silicon

through full scale 3D device simulation. Very good agreement with experimental mobility values up to concentrations of $1 \times 10^{18} \text{ cm}^{-3}$ was achieved. Information regarding all aspects of the simulation was given in support of the accuracy of such a claim. Some artificial carrier heating was observed at the highest doping concentrations but is attributed to the sampling of the electron energies during more frequent short range interactions, rather than inaccurate integration of the equations of motion. Comparison of the simulation of the continuous N-I-N diode in Chapter 3 and the simulation of an atomistically doped copy of the same device showed excellent agreement, further validating the “atomistic” simulation approach. The influence of the random discrete dopants was clearly seen producing variation within the depletion regions, and non-uniform energy and velocity distributions within the intrinsic region, as compared with the continuous simulation.

6.3 Transport Variation

The *ab initio* approach of introducing Coulomb scattering in the MC simulations was then applied to investigate the transport variations associated with the random number and position of dopant atoms within a device. This necessitated the development of a charge assignment scheme in the continuously doped region which reproduces the effect of CIC charge assignment in the random dopant region. Such a scheme guarantees consistency in simulation regions where the doping description changes from continuous to atomistic. It is generally applicable within continuous device simulation and provides a better balance between the electron concentration and the doping concentration where in the MC simulations the electrons are assigned via the CIC scheme.

Initially, the effect on the transport of a single trapped charge within an otherwise continuously doped MOSFET was investigated. The MC simulations for devices with square channels of length 30 nm , 20 nm and 10 nm were compared to DD simulation. As stated

previously, the DD simulations only capture the electrostatic effects while MC including *ab initio* scattering properly captures additional transport variations. The simulations are carried out at low drain bias to allow fair comparison of DD and MC results by minimizing non equilibrium transport. The DD mobility models are calibrated to match bulk MC simulation results. MC simulation is performed in the frozen field approximation where the steady-state potential and initial electron concentration are imported from a DD solution. This is justifiable at the low drain bias and significantly speeds up the simulation. Comparison of the MC and DD results provides information on the relative importance of the electrostatic and transport effects. In all devices, over a range of gate voltages, an increase in the current reduction is observed upon introducing a trapped charge within MC simulation. The additional effect associated with the transport variation due to scattering from the trapped charge is most significant at high gate voltages. This implies that at low gate voltages the electrostatic impact of the trapped charge in reducing the number of carriers within the channel is most significant, while at high gate voltages additional scattering from the trapped charge dominates the reduction in current. The same trend was reproduced in successively smaller devices, but in each case the percentage current reduction, and therefore the contribution from additional scattering, increased.

Following the illustrative example of a single trapped charge, an ensemble of 50 30 nm gate length MOSFETs with fully atomistic doped channels was simulated under the same low drain bias in the frozen field approximation. The comparison of the results from DD and MC simulation show a significant increase in the on current variation associated with additional transport variation. The correlation of the results from DD and MC simulation again implies the importance of the donor positions within the channel. In particular, three devices were highlighted in which the percentage variation in current associated with the atomistic nature of the channel doping were comparable in the DD results, but show large differences in the MC simulation.

6.4 Future Work

The *ab initio* approach to Coulomb scattering described in this work has had success in describing random dopant effects and implies the possible benefit of adopting similar *ab initio* approaches to other scattering mechanisms. In particular, surface roughness scattering has so far been omitted entirely from this work but may be directly incorporated through the real space scattering from a randomly generated rough interface surface. It remains to be seen to what extent such an approach is capable of reproducing the universal mobility curve in silicon, but such a model could readily be incorporated into the existing 3D *ab initio* EMC simulator to study transport variations associated with the random pattern of the interface in individual transistors. Similarly, the effects of variation in the high- κ gate stack on the surface potential will introduce position dependent scattering and transport variation which could be addressed in a similar manner.

Fully self consistent EMC simulations of random dopant effects in MOSFET devices need to be performed to evaluate the impact of scattering at high drain voltage. In such an operating condition, the Coulomb scattering is expected to play a less significant role in the high field region near the drain which may somewhat reduce the variations at high drain bias conditions. Such self consistent simulations are at present very time consuming and may require the use of non-uniform meshing to improve efficiency. Similarly, the investigation of more sophisticated integration routines with better energy conserving properties may allow the limitation on the field adjusting time step to be relaxed, reducing the number of simulation time steps required per simulation and reducing the overall simulation time.

Bibliography

- [1] “International Technology Roadmap for Semiconductors 2004 update edition,” 2004.
- [2] W. M. Siu, “Foreword,” *Intel Technology Journal*, vol. 8, 2004.
- [3] R. Chau, J. Kavalieros, B. Roberds, R. Schenker, D. Lionberger, D. Barlage, B. Doyle, R. Arghavani, A. Murthy, and G. Dewey, “30nm Physical Gate Length CMOS Transistor with 1.0 ps *n*-MOS and 1.7 ps *p*-MOS Gate Delays,” *IEDM Tech. Dig. 2000*, pp. 45–48, 2000.
- [4] S. Inaba, K. Okano, S. Matsuda, M. Fujiwara, A. Hokazono, K. Adachi, K. Ohuchi, H. Suto, H. Fukui, T. Shimizu, S. Mori, H. Oguma, A. Murakoshi, T. Itani, M. Takayanagi, A. Azuma, H. Oyamatsu, K. Suguro, Y. Katsumata, Y. Toyoshima, and H. Ishiuchi, “High Performance 35nm Gate Length CMOS with *NO* Oxynitride Gate Dielectric and *Ni* SALICIDE,” *IEDM Tech. Dig. 2001*, pp. 641–644, 2001.
- [5] R. Chau, “30nm and 20nm Physical Gate Length CMOS Transistors,” *Proc. Silicon Nanoelectronics Workshop 2001*, 2001.
- [6] F. Boeuf, T. Skotnicki, S. Monfray, C. Julien, D. Dutartre, J. Mazoyer, R. Palla, B. Tavel, P. Ribot, E. Sondergard, and M. Sanquer, “16nm planar *n*MOSFET man-

BIBLIOGRAPHY

- ufacturable within state-of-the-art CMOS process thanks to specific design and optimisation,” *IEDM Tech. Dig. 2001*, pp. 637–640, 2001.
- [7] B. Yu, H. Wang, A. Joshi, Q. Xiang, E. Ibok, and M. Lin, “15 nm Gate Length Planar CMOS Transistor,” *IEDM Tech. Dig. 2001*, pp. 937–939, 2001.
- [8] A. Hokazono, K. Ohuchi, M. Takayanagi, Y. Watanabe, S. Magoshi, Y. Kato, T. Shimizu, S. Mori, H. Oguma, T. Sasaki, H. Yoshimura, K. Miyano, N. Yasutake, H. Suto, K. Adachi, H. Fukui, T. Watanabe, N. Tamaoki, Y. Toyoshima, and H. Ishiuchi, “14 nm Gate Length CMOSFETs Utilizing Low Thermal Budget Process with Poly-SiGe and Ni Salicide,” *IEDM Tech. Dig. 2002*, pp. 639–642, 2002.
- [9] H. Wakabayashi, S. Yamagami, N. Ikezawa, A. Ogura, M. Narihiro, K. Arai, Y. Ochiai, K. Takeuchi, T. Yamamoto, and T. Mogami, “Sub-10-nm Planar-Bulk-CMOS Devices using Lateral Junction Control,” *IEDM Tech. Dig. 2003*, pp. 989–991, 2003.
- [10] B. Yu, L. Chang, S. Ahmed, H. Wang, S. Bell, C. Yang, C. Tabery, C. Ho, Q. Xiang, T. King, J. Bokor, C. Hu, M. Lin, and D. Kyser, “FinFET Scaling to 10 nm Gate Length,” *IEDM Tech. DIG 2002.*, pp. 251–254, 2002.
- [11] B. Doris, M. Leong, T. Kamarsky, Y. Zhang, R. A. Roy, O. Dokumachi, Z. Ren, F. Jamin, L. Shi, W. Natzle, H. Huang, J. Mazzapella, A. Mocuta, S. Womack, M. Gribelyuk, E. C. Jones, r. J. Miller, H.-S. P. Wong, and W. Haensch, “Extreme Scaling with Ultra-Thin Si Channel MOSFETs,” *IEDM Tech. Dig. 2002*, pp. 267–270, 2002.
- [12] K. A. Bowman, X. Tang, J. C. Eble, and J. D. Meindl, “Impact of Extrinsic and Intrinsic Parameter Fluctuations on CMOS Circuit Performance,” *IEEE J. Solid-State Circuits*, vol. 35, pp. 1186–1193, 2000.

BIBLIOGRAPHY

- [13] A. J. Bhavnagarwala, X. Tang, and J. D. Meindl, "The Impact of Intrinsic Device Fluctuations on CMOS SRAM Cell Stability," *IEEE J. Solid-State Circuits*, vol. 36, pp. 658–665, 2001.
- [14] X. Tang, V. K. De, and J. D. Meindl, "Intrinsic MOSFET Parameter Fluctuations Due to Random Dopant Placement," *IEEE Trans. VLSI Systems*, vol. 5, pp. 369–376, 1997.
- [15] S. Roy, B. Cheng, and A. Asenov, "A methodology for quantitatively introducing 'atomistic' fluctuations in compact device models for circuit analysis," *J. Comp. Elec.*, vol. 2, pp. 427–431, 2003.
- [16] B. J. Cheng, S. Roy, G. Roy, and A. Asenov, "Integrating 'atomistic' intrinsic parameter fluctuations into compact model circuit analysis," *Proc. ESSDERC 2003*, pp. 436–440, 2003.
- [17] B. J. Cheng, S. Roy, G. Roy, F. Adamu-Lema, and A. Asenov, "The Impact of Random Doping Effects on Decanano CMOS SRAM Cell Stability," *Proc. ESSCIRC 2004*, pp. 219–222, 2004.
- [18] G. P. Patsis, V. Constantoudis, A. Tserepi, and E. Gogolides, "Roughness analysis of lithographically produced nanostructures: off line measurement, scaling analysis and Monte Carlo simulations," *Microelectronic Engineering*, vol. 67, pp. 319–325, 2003.
- [19] H. Kim, J. Lee, J. Shin, S. Woo, H. Cho, and J. Moon, "Experimental Investigation of the Impact of LWR on Sub-100-nm Device Performance," *IEEE Trans. Elec. Dev.*, vol. 51, pp. 1984–1988, 2004.

BIBLIOGRAPHY

- [20] A. Asenov, S. Kaya, and A. R. Brown, "Intrinsic Parameter Fluctuations in Decananometer MOSFETs Introduced by Gate Line Edge Roughness," *IEEE Trans. Elec. Dev.*, vol. 50, pp. 1254–1260, 2003.
- [21] G. Roy, *Simulation of Intrinsic Parameter Fluctuations in Nano CMOS Devices*. PhD thesis, Dept. Electronics & Electrical Engineering University of Glasgow, 2005.
- [22] B. Hoeneisen and C. A. Mead, "Fundamental limitations in micro-electronics –I MOS technology," *Solid State Electron.*, vol. 15, pp. 819–829, 1972.
- [23] R. W. Keyes, "Physical limits in digital electronics," *Proc. IEEE*, vol. 63, pp. 740–766, 1975.
- [24] A. Asenov, "Random Dopant Induced Threshold Voltage Lowering and Fluctuations in Sub-0.1 μm MOSFETs: A 3-D 'Atomistic' Simulation Study," *IEEE Trans. Elec. Dev.*, vol. 45, pp. 2505–2510, 1998.
- [25] A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, "Simulation of Intrinsic Parameter Fluctuations in Decananometer and Nanometer-Scale MOSFETs," *IEEE Trans. Elec. Dev.*, vol. 50, pp. 1837–1852, 2003.
- [26] F. Adamu-Lema, *Scaling and Intrinsic Parameter Fluctuations in nano CMOS Devices*. PhD thesis, Dept. Electronics & Electrical Engineering University of Glasgow, 2005.
- [27] "International Technology Roadmap for Semiconductors 2002 update edition," 2002.

BIBLIOGRAPHY

- [28] R. W. Keyes, "Effect of Randomness in the Distribution of Impurity Ions on FET Threshold in Integrated Electronics," *IEEE J. Solid-State Circuits*, pp. 245–247, August 1975.
- [29] K. R. Lakshmikumar, R. A. Hadaway, and M. A. Copeland, "Characterization and Modeling of Mismatch in MOS Transistors for Precision Analog Design," *IEEE J. Solid-State Circuits*, vol. 21, pp. 1057–1066, 1986.
- [30] T. Mikolajick and H. Ryssel, "Influence of Statistical Dopant Fluctuations on MOS Transistors with Deep Submicron Channel Lengths," *Microelectron. Eng.*, vol. 21, pp. 419–422, 1993.
- [31] D. Burnett, K. Erington, C. Subramanian, and K. Baker, "Implications of fundamental threshold voltage variations for high density SRAM and logic circuits," *VLSI Symp. Tech. Dig. 1994*, pp. 15–16, 1994.
- [32] M. Eisele, J. Berthold, R. Thewes, E. Wohlrabh, D. Schmitt-Landsiedel, and W. Weber, "Intra-die device parameter variation and their impact on digital CMOS gates at low supply voltages," *IEDM Tech. Dig. 1995*, pp. 67–70, 1995.
- [33] H. P. Tuinhout, "Impact of parametric mismatch and fluctuations on performance and yield of deep-submicron CMOS technologies," *Proc. ESSDERC 2002*, pp. 95–101, 2002.
- [34] J. Shyu, G. C. Temes, and F. Krummenacher, "Random Error Effects in Matched MOS Capacitors and Current Sources," *IEEE J. Solid-State Circuits*, vol. 19, pp. 948–955, 1984.

BIBLIOGRAPHY

- [35] M. Steyaert, J. Bastos, R. Roovers, O. Kinget, W. Sansen, B. Graindourse, A. Pergot, and E. Janssens, "Threshold voltage mismatch in short-channel MOS transistors," *Electron. Lett.*, vol. 30, pp. 1546–1548, 1994.
- [36] R. Difrenza, P. Llinares, and G. Gibaudo, "Impact of short channel and quantum effects on the MOS transistor mismatch," *Proc. 3rd European Workshop on Ultimate Integration of Silicon 2002*, pp. 127–130, 2002.
- [37] O. R. dit Buisson and G. Morin, "MOSFET matching in deep submicron technology," *Proc. ESSDERC 1996*, pp. 731–734, 1996.
- [38] J. T. Hortsman, U. Hilleringmann, and K. F. Gosser, "Matching analysis of deposition defined 50-nm MOSFETs," *IEEE Trans. Electron. Dev.*, vol. 45, pp. 299–306, 1998.
- [39] T. Mizuno, J. Okamura, and A. Toriumi, "Experimental Study of Threshold Voltage Fluctuations Using An 8k MOSFET's Array," *VLSI Tech. Dig. 1993*, pp. 41–42, 1993.
- [40] T. Mizuno, M. Iwase, H. Niiyama, T. Shibata, K. Fujisaki, T. Nakasugi, A. Toriumi, and U. Ushiku, "Performance FLuctuations of 0.1 μm MSOFETs - Limitation of 0.1 μm ULSIs," *Proc. VLSI Symp. 1994*, pp. 13–14, 1994.
- [41] T. Mizuno, J. Okamura, and A. Toriumi, "Experimental Study of Threshold Voltage Fluctuation Due to Statistical Variation of Channel Dopant Number in MOSFET's," *IEEE Trans. Elec. Dev.*, vol. 41, pp. 2216–2221, 1994.
- [42] T. Mizuno, "Influence of Statistical Spatial-Nonuniformity of Dopant Atoms on Threshold Voltage in a System of Many MOSFETs," *Jpn. J. Appl. Phys.*, vol. 35, pp. 842–848, 1996.

BIBLIOGRAPHY

- [43] T. Hagivaga, K. Yamaguchi, and S. Asai, "Threshold Voltage Variation in Very Small MOS Transistors due to Local Dopant Fluctuations," *Proc. Symp. VLSI Technol. Tech. Dig. 1982*, pp. 46–47, 1982.
- [44] K. Takeuchi, T. Tatsumi, and A. Furukawa, "Channel Engineering for the Reduction of Random-Dopant-Placement-Induced Threshold Voltage Fluctuations," *IEDM Tech. Dig. 1996*, pp. 841–844, 1996.
- [45] Y. Taur, D. A. Buchanan, W. Chen, D. J. Frank, K. E. Ismail, S. Lo, A. A. Sai-Hakasz, R. G. Viswanathan, H. J. C. Wann, S. J. Wind, and H. Wong, "CMOS scaling into the nanometer regime," *Proc. IEEE*, vol. 85, pp. 486–504, 1997.
- [46] P. A. Stolk, F. P. Widdershoven, and D. B. M. Klassen, "Modelling Statistical Dopant Fluctuations in MOS Transistors," *IEEE Trans. Elec. Dev.*, vol. 45, pp. 1960–1971, 1998.
- [47] K. Nishinohara, N. Shigyo, and T. Wada, "Effects of Microscopic Fluctuations in Dopant Distributions on MOSFET Threshold Voltage," *IEEE Trans. Elec. Dev.*, vol. 39, pp. 634–639, 1992.
- [48] P. A. Stolk and D. B. M. Klassen, "The Effect of Statistical Dopant Fluctuations on MOS Device Performance," *IEDM Tech. Dig. 1996*, pp. 627–630, 1996.
- [49] H. Wong and Y. Taur, "Three-Dimensional 'Atomistic' Simulation of Discrete Random Dopant Distribution Effects in Sub- $0.1\mu\text{m}$ MOSFET's," *Proc. IEDM 93*, pp. 705–708, 1993.
- [50] D. J. Frank, Y. Taur, M. Jeong, and H. P. Wong, "Monte Carlo Modeling of Threshold Voltage Variation due to Dopant Fluctuations," *Symp. VLSI Circuits Tech. Dig. 1999*, pp. 171–172, 1999.

BIBLIOGRAPHY

- [51] M. Jaraiz, G. H. Gilmer, and J. M. Poate, "Atomistic calculations of ion implantation in Si: Point defect and transient enhanced diffusion phenomena," *Appl. Phys. Lett.*, vol. 68, pp. 409–411, 1996.
- [52] M. Jaraiz, P. Castrillo, R. Pinacho, I. Martin-Bragado, and J. Barbolla, "Atomistic Front-End Process Modelling: A Powerful Tool for Deep-Submicron Device Fabrication," *Proc. SISPAD 2001*, pp. 10–18, 2001.
- [53] A. Asenov, M. Jaraiz, S. roy, G. Roy, F. Adamu-Lema, A. R. Brown, V. Moroz, and R. Gafiteanu, "Integrated atomistic process and device simulation of decananometer MOSFETs," *Proc. SISPAD 2002*, pp. 87–90, 2002.
- [54] T. Ezaki, T. Ikezava, A. Notsu, K. Tanaka, and M. Hane, "3D MOSFET Simulation Considering Long-Range Coulomb Potential Effects for Analysing Statistical Dopant Induced Fluctuations Associated with Atomistic Process Simulation," *Proc SISPAD 2002*, pp. 91–94, 2002.
- [55] T. Ezaki, T. Ikezawa, and M. Hane, "Investigation of Realistic Dopant Fluctuation Induced Device Characteristics Variation for Sub-100nm CMOS by Using Atomistic 3D Process/Device Simulator," *Proc. IEDM 2002*, pp. 311–314, 2002.
- [56] A. Asenov and S. Saini, "Suppression of Random Dopant Induced Threshold Voltage Fluctuations in sub-0.1 μm MOSFETs with Epitaxial and δ -doped Channels," *IEEE Trans. Elec. Dev.*, vol. 46, pp. 1718–1723, 1999.
- [57] A. Asenov, R. Balasubramaniam, A. R. Brown, and J. H. Davies, "Hierarchical Approach to 'atomistic' 3D MOSFET Simulation," *IEEE Trans. Comput. Aided Des. Integrated Circuits Syst.*, vol. 18, pp. 1558–1565, 1999.

BIBLIOGRAPHY

- [58] N. Sano and M. Tomizawa, "Random dopant model for three-dimensional drift-diffusion simulations in metal-oxide-semiconductor-field-effect-transistors," *Appl. Phys. Lett.*, vol. 79, pp. 2267–2269, 2001.
- [59] G. Roy, A. R. Brown, A. Asenov, and S. Roy, "Quantum Aspects of Resolving Discrete Charges in 'Atomistic' Device Simulations," *J. Comp. Elec.*, vol. 2, pp. 323–327, 2003.
- [60] G. Roy, A. R. Brown, A. Asenov, and S. Roy, "Bipolar Quantum Corrections in Resolving Individual Dopants in 'Atomistic' Device Simulation," *Superlattices and Microstructures*, vol. 34, pp. 327–334, 2003.
- [61] M. G. Ancona and G. J. Iafrate, "Quantum correction to the Equation of State of an Electron Gas in a Semiconductor," *Phys. Rev. B*, vol. 39, pp. 9536–9540, 1989.
- [62] A. Asenov, J. R. Watling, A. R. Brown, and D. K. Ferry, "The Use of Quantum Potentials for Confinement and Tunnelling in Semiconductor Devices," *J. Comp. Elec.*, vol. 1, pp. 503–513, 2002.
- [63] A. Asenov, A. R. Brown, and J. R. Watling, "Quantum corections in the simulation of decanano MOSFETs," *Solid State Elec.*, vol. 47, pp. 1141–1145, 2003.
- [64] A. Avellan, W. Krautschneider, and S. Schwantes, "Observations and Modeling of Random Dopant Telegraph Signals in the Gate and Drain Currents of Tunneling Metal-Oxide-Semiconductor-Field-Effect-Transistors," *Appl. Phys. Lett.*, vol. 78, pp. 2790–2792, 2001.
- [65] A. Asenov, R. Balasubramaniam, A. R. Brown, and J. H. Davies, "RTS Amplitudes in Decananometer MOSFETs: a 3D Simulation Study," *IEEE Trans. Elec. Dev.*, vol. 50, pp. 839–845, 2003.

BIBLIOGRAPHY

- [66] A. Asenov, R. Balasubramaniam, A. R. Brown, J. H. Davies, and S. Saini, "Random Telegraph Signal Amplitudes in Sub 100nm (decanano) MOSFETs: a 3D 'atomistic' Simulation Study," *IEDM Tech. Dig. 2000*, pp. 279–282, 2000.
- [67] H. M. Bu, Y. Shi, X. L. Yuan, Y. D. Zhend, S. H. Gu, H. Majima, H. Ishicuro, and T. Hiramoto, "Impact of the Device Scaling on the Low-Frequency Noise in *n*-MOSFETs," *Appl. Phys. A*, vol. 71, pp. 133–136, 2000.
- [68] A. Godoy, F. Gamiz, A. Palma, J. A. Jimenez-Tejada, J. Banqueri, and J. A. Lopez-Villanueva, "Influence of mobility fluctuations on random telegraph signal amplitude in *n*-channel metal-oxide-semiconductor-field-effect-transistors," *J. Appl. Phys.*, vol. 82, pp. 4621–4628, 1997.
- [69] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. Springer-Verlag Wien New York, 1984.
- [70] D. M. Caughy and R. E. Thomas, "Carrier Mobilities in Silicon Empirically Related to Doping and Field," *Proc. IEEE*, vol. 52, pp. 2192–2193, 1967.
- [71] T. Grasser, T. Tang, H. Kosina, and S. Selberherr, "A review of Hydrodynamic and Energy Transport Models for Semiconductor Device Simulation," *Proc. IEEE*, vol. 91, pp. 251–274, 2003.
- [72] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng, "Random Telegraph Noise of Deep-Submicrometer MOSFETs," *IEEE Elec. Dev. Lett.*, vol. 11, pp. 90–92, 1990.
- [73] Z. Shi, J. P. Mieville, and M. Dutoit, "Random telegraph signals in deep submicron *n*-MOSFETs," *IEEE Trans. Elec. Dev.*, vol. 41, pp. 1161–1168, 1994.

BIBLIOGRAPHY

- [74] S. T. Martin, G. P. Li, E. Worley, and J. White, "The gate bias and geometry dependence of random telegraph signal amplitudes," *IEEE Elec. Dev. Lett.*, vol. 18, pp. 444–446, 1997.
- [75] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*. Springer-Verlag Wien New York, 1989.
- [76] R. P. Joshi and D. K. Ferry, "Effect of multi-ion screening on the electronic transport in deoped semiconductors: A molecular dynamics analysis," *Phys. Rev. B*, vol. 43, pp. 9734–9739, 1991.
- [77] W. J. Gross, D. Vasileska, and D. K. Ferry, "A Novel Approach for Introducing the Electron-Electron and Electron-Impurity Interactions in Particle Based Simulations," *IEEE Trans. Elec. Dev. Lett.*, vol. 20, pp. 463–465, 1999.
- [78] D. Vasileska, W. J. Gross, and D. K. Ferry, "Monte Carlo particle-based simulations of deep-submicron n-MOSFETs with real-space treatment of electron-electron and electron-impurity interactions," *Superlattices and Microstructures*, vol. 27, pp. 147–157, 2000.
- [79] C. J. Wordelman and U. Ravaioli, "Integration of a Particle-Particle-Particle-Mesh Algorithm with the Ensemble Monte Carlo Method for the Simulation of Ultra-Small Semiconductor Devices," *IEEE Trans. Elec. Dev.*, vol. 47, pp. 410–416, 2000.
- [80] W. J. Gross, D. Vasileska, and D. K. Ferry, "Ultrasmall MOSFETs: The Importance of the Full Coulomb Interaction on Device Characteristics," *IEEE Trans. Elec. Dev.*, vol. 47, pp. 1831–1837, 2000.

BIBLIOGRAPHY

- [81] W. J. Gross, D. Vasileska, and D. K. Ferry, "Three-dimensional Simulations of ultra small metal-oxide-semiconductor-field-effect-transistors: The role of the discrete impurities on the device terminal characteristics," *J. Appl. Phys.*, vol. 91, pp. 3737–3740, 2002.
- [82] 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 Trans. Nanotech.*, vol. 4, pp. 195–200, 2002.
- [83] M. Nagase, H. Namatsu, K. Kurihara, K. Iwadate, K. Murase, and T. Makino, "Nano-scale fluctuations in electron beam resist pattern evaluated by atomic force microscopy," *Microelectron. Eng.*, vol. 30, pp. 419–422, 1996.
- [84] H. Namatsu, M. Nagase, T. Yamaguchi, K. Yamazaki, and K. Kurihara, "Influence of edge roughness in resist patterns on etched patterns," *J. Vac. Sci. Technol. B*, vol. 16, pp. 3315–3321, 1998.
- [85] T. Yoshimura, H. Shairaishi, J. Yamamoto, and S. Okazaki, "Nano edge roughness in polymer resist patterns," *Appl. Phys. Lett.*, vol. 63, pp. 764–766, 1993.
- [86] D. A. Buchanan, "Scaling the gate dielectric: Materials, integration & reliability," *IBM J. Res. Dev.*, vol. 43, pp. 245–264, 1999.
- [87] M. Niwa, T. Kouzaki, K. Okada, M. Udagawa, and R. Sinclair, "Atomic-Order Planarization of Ultrathin SiO₂/Si(001) Interfaces," *Jpn. J. Appl. Phys.*, vol. 33, pp. 388–394, 1994.
- [88] S. M. goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, "Surface Roughness at the Si(100)-SiO₂ Interface," *Phys. Rev. B*, vol. 32, pp. 8171–8186, 1985.

BIBLIOGRAPHY

- [89] A. Asenov and S. Kaya, "Effect of Oxide Roughness on the Threshold Voltage Fluctuations in decanano MOSFETS with Ultrathin Gate Oxide," *Proc. SISPAD 2000*, pp. 135–138, 2000.
- [90] A. Asenov, S. Kaya, and J. H. Davies, "Intrinsic Threshold Voltage Fluctuations in decanano MOSFETs due to Local Oxide Thickness Variations," *IEEE Trans. Elec. Dev.*, vol. 49, pp. 112–119, 2002.
- [91] C. Riddet, A. R. 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," *Proc. IWCE 10*, 2004.
- [92] C. Moglestue, "A Self-Consistent Monte Carlo Particle Model to Analyze Semiconductor Microcomponents of any Geometry," *IEEE Trans. Computer-Aided Design*, vol. 5, pp. 326–345, 1986.
- [93] E. Sangiorgi, B. Ricco, and F. Venturi, "MOS²: An Efficient MOnTe Carlo Simulator for MOS Devices," *IEEE Trans. Computer-Aided Design*, vol. 7, pp. 259–271, 1988.
- [94] F. Venturi, R. K. Smith, E. Sangiorgi, M. R. Pinto, and B. Ricco, "A General Purpose Device Simulator Coupling Poisson and Monte Carlo Transport Applications to Deep Submicron MOSFETs," *IEEE Trans. Computer-Aided Design*, vol. 8, pp. 360–369, 1989.
- [95] P. Lugli, "The MOnTe Carlo Method for Semiconductor Device and Process Modeling," *IEEE Trans. Computer-Aided Design*, vol. 9, pp. 1164–1176, 1990.

BIBLIOGRAPHY

- [96] H. Kosina, M. Nedjalkov, and S. Selberherr, "Theory of the Monte Carlo Method for Semiconductor Device Simulation," *IEEE Trans. Elec. Dev.*, vol. 47, pp. 1898–1908, 2000.
- [97] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Brooks/Cole, 1976.
- [98] H. E. Hall, *Solid State Physics*. London: Wiley, 1974.
- [99] K. Tomizawa, *Numerical Simulation of Submicron Semiconductor Devices*. Artech House, 1993.
- [100] Y. Taur and T. Ning, *Fundamentals of Modern VLSI Devices*. Cambridge University Press, 1998.
- [101] F. Rossi, P. Poli, and C. Jacoboni, "Weighted Monte Carlo approach to electron transport in semiconductors," *Semicond. Sci. Technol.*, vol. 7, pp. 1017–1035, 1992.
- [102] D. K. Ferry, *Semiconductor Transport*. Taylor & Francis, 2000.
- [103] J. R. Chelikowsky and M. L. Cohen, "Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors," *Phys. Rev. B*, vol. 14, pp. 556–5582, 1976.
- [104] V. Romano, "Non-parabolic band hydrodynamical model of silicon semiconductors and simulation of electron devices," *Math. Meth. Appl. Sci.*, vol. 24, no. 439–471, 2001.
- [105] C. Moglestue, *Monte Carlo Simulation of Semiconductor Devices*. Chapman & Hall, 1993.

BIBLIOGRAPHY

- [106] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. Alerigi-Quaranta, “Electron Drift Velocity in Silicon,” *Phys. Rev. B*, vol. 12, pp. 2265–2284, 1975.
- [107] M. V. Fischetti, “Monte Carlo Simulation of Transport in Technologically Significant Semiconductors of the Diamond and Zinc-Blende Structures - Part I: Homogeneous Transport,” *IEEE Trans. Elec. Dev.*, vol. 38, pp. 634–649, 1991.
- [108] M. V. Fischetti and S. E. Laux, “Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space charge effects,” *Phys. Rev. B*, vol. 38, pp. 9721–9745, 1988.
- [109] F. M. Bufler, *Full-Band Monte Carlo Simulation of Electrons and Holes in Strained Si and SiGe*. Herbert Utz Verlag, 1998.
- [110] R. Clerc, , F. Dauge, R. Hude, and G. Ghibaudo, “Quantum Confinement in NMOS Double Gate Devices: Impact of Conduction Band Non Parabolicity,” *Proc. ULIS 2004*, 2004.
- [111] M. Lundstrom, *Fundamentals of Carrier Transport*. Cambridge University Press, 2nd edition ed., 2000.
- [112] H. P. Myres, *Introductory Solid State Physics*. Taylor & Francis, 2nd edition ed., 1997.
- [113] H. D. Rees, “Calculation of Steady State Distribution Function by Exploiting Stability,” *Phys. Lett. A*, vol. 26, pp. 4126–4127, 1968.
- [114] C. Jacoboni and L. Reggiani, “The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials,” *Rev. Mod. Phys.*, vol. 55, pp. 645–705, 1983.

BIBLIOGRAPHY

- [115] P. J. Price, "The Theory of Hot Electrons," *IBM J. Res. Dev.*, vol. 14, pp. 12–24, 1970.
- [116] X. Zhou, R. J. Bowman, and T. Y. Hsiang, "EMCUR: An Ensemble Monte Carlo Program for III-V Compound Semiconductor Device Modeling and Simulation," tech. rep., Nanyang Technological University School of Electrical and Electronic Engineering Report, 1990.
- [117] W. R. Thurber, R. L. Mattis, and Y. M. Liu, "Resistivity-Dopant density Relationship for Phosphorus Doped Silicon," *J. Electrochem. Soc.: Solid State Sci. & Technol.*, vol. 127, pp. 1807–1812, 1980.
- [118] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles*. New York: McGraw-Hill, 1981.
- [119] C. Millar, A. Asenov, and J. R. Watling, "Excessive Over-Relaxation Method for Multigrid Poisson Solvers," *J. Comp. Elec.*, vol. 1, pp. 341–345, 2002.
- [120] S. E. Laux, "On particle-mesh coupling in Monte Carlo semiconductor device simulation," *IEEE Trans. Computer-Aided Design*, vol. 15, pp. 1266–1277, 1996.
- [121] P. W. Rambo and J. Denavit, "Time Stability of Monte Carlo Device Simulation," *IEEE Trans. Computer-Aided Design of Integrated Circuits*, vol. 12, pp. 1734–1741, 1993.
- [122] R. Clerc, P. Palestri, and A. Abramo, "Investigation of Convergence and Stability of Self-Consistent Monte Carlo Device Simulations," *Proc. ESSDERC 2002*, pp. 191–194, 2002.
- [123] T. Gonzalez and D. Pardo, "Physical Models of Ohmic Contact for Monte Carlo Device Simulation," *Solid State Electron.*, vol. 39, pp. 555–562, 1996.

BIBLIOGRAPHY

- [124] D. L. Woolard, H. Tian, M. A. Littlejohn, and K. W. Kim, "The Implementation of Physical Boundary Conditions in the Monte Carlo Simulation of Electron Devices," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 13, pp. 1241–1246, 1994.
- [125] P. D. Yoder, K. Gartner, and W. Fichtner, "A Generalized Ramo-Shockley Theorem for Classical to Quantum Transport at Arbitrary Frequencies," *J. Appl. Phys.*, vol. 79, pp. 1951–1959, 1996.
- [126] J. Zimmermann and E. Constant, "Application of Monte Carlo techniques to hot carrier diffusion noise calculation in unipolar semiconducting components," *Solid State Electron.*, vol. 23, pp. 915–925, 1980.
- [127] D. Chattopadhyay and H. J. Queisser, "Electron scattering by ionized impurities in semiconductors," *Rev. Mod. Phys.*, vol. 53, pp. 745–768, 1981.
- [128] B. K. Ridley, *Quantum Processes in Semiconductors*. Oxford University Press, 4th edition ed., 1999.
- [129] A. Messiah, *Quantum Mechanics*. North Holland Publishing Co., 1961.
- [130] J. R. Watling, L. Yang, M. Borici, J. R. Barker, and A. Asenov, "Degeneracy and high Doping Effects in Sub-Micron Relaxed and Strained Si *n*-MOSFETs," *J. Comp. Elec.*, vol. 2, pp. 475–479, 2003.
- [131] N. Takemoto, "On the Screening of Impurity Potential by Conduction Electrons," *J. Phys. Soc. Jpn.*, vol. 14, pp. 1142–1158, 1959.
- [132] P. Davies and P. L. Giles, "Compensation in heavily doped *n*-type InP and GaAs," *J. Appl. Phys.*, vol. 58, pp. 3059–3067, 1985.

BIBLIOGRAPHY

- [133] B. K. Ridley, "Reconciliation of the Conwell-Weisskopf and Brooks-Herring formulae for charged-impurity scattering in semiconductors: Third-body interference," *J. Phys. C*, vol. 10, pp. 1589–1593, 1977.
- [134] C. K. Birdsall and A. B. Langdon, *Plasma Physics Via Computer Simulation*. Bristol: Hilger, 1991.
- [135] W. H. Press, S. A. Teuolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*.
- [136] A. Leach, *Molecular Modelling Principles and Applications*. Pearson Prentice Hall, 2001.
- [137] S. P. Goel and J. W. Eastwood, "A 10 000 Particle Molecular Dynamics Model with Long-Range Forces," *Chem. Phys. Lett.*, vol. 21, pp. 589–591, 1973.
- [138] J. W. Eastwood and R. W. Hockney, "Shaping the Force Law in Two-Dimensional Particle-Mesh Models," *J. Comput. Phys.*, vol. 16, pp. 342–259, 1974.
- [139] S. M. Ramey and D. K. Ferry, "A New Model for Including Discrete Dopant Ions into Monte Carlo Simulations," *IEEE Trans. Nano.*, vol. 2, pp. 193–197, 2003.
- [140] C. R. Arokianathan, J. H. Davies, and A. Asenov, "*ab initio* Coulomb Scattering in Atomistic Device Simulation," *VLSI Design*, vol. 8, pp. 331–335, 1998.
- [141] C. R. Arokianathan, *Methods for the Atomistic Simulation of Ultrasmall Semiconductor Devices*. PhD thesis.
- [142] S. Barraud, P. Dollfus, S. Galdin, and P. Hesto, "Short-range and long-range Coulomb interactions for 3D Monte Carlo device simulation with discrete impurity distribution," *Solid-State Electron.*, vol. 46, pp. 1061–1067, 2002.

BIBLIOGRAPHY

- [143] S. Barraud, P. Dollfus, S. Galdin, and P. Hesto, "Effects of random discrete impurities in ultra-short MOSFET using 3D Monte Carlo simulation," *Proc. ESSDERC 2002*, 2002.
- [144] P. Dollfus, A. Bournel, S. Galdin, S. Barraud, and P. Hesto, "Effect of Discrete Impurities on Electron Transport in Ultrashort MOSFET Using 3D MC Simulation," *IEEE Trans. Elec. Dev.*, vol. 51, pp. 749–756, 2004.
- [145] K. Natori, T. Shimizu, and T. Ikenobi, "Multi Subband Effects on Performance Limit of Nanoscale MOSFETs," *Jpn. J. Appl. Phys.*, vol. 42, pp. 2063–2066, 2003.
- [146] E. Simoen, A. Mercha, L. Pantisano, and C. Claeys, "Tunnelling $1/f$ noise in 5nm HfO_2 / 2.1 nm SiO_2 Gate stack n -MOSFETs," *Proc. 5th European Workshop on Ultimate Integration of Silicon 2004*, pp. 31–34, 2004.
- [147] K. A. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant, "Discrete resistance switching in submicron silicon inversion layers: Individual interface traps and low frequency ($1/f$) noise," *Phys. Rev. Lett.*, vol. 52, pp. 228–231, 1984.
- [148] M. J. Kirton and M. J. Uren, "Noise in solid state microstructures: A new perspective on individual defects, interface states and low frequency ($1/f$) noise," *Adv. Phys.*, vol. 38, pp. 367–368, 1989.
- [149] P. J. Restle, J. W. Park, and B. F. Lloyd, "DRAM variable retention time," *IEDM Tech. Dig.*, pp. 807–810, 1992.
- [150] A. Asenov, R. Balasubramaniam, A. R. Brown, and J. H. Davies, "Effect of single-electron interface trapping in decanano MOSFETs: a 3D atomistic simulation study," *Superlattices and Microstructures*, vol. 27, pp. 411–416, 2000.

BIBLIOGRAPHY

- [151] K. Hess, A. Haggag, W. McMahon, B. Fischer, K. Cheng, J. Lee, and J. Lyding, "Simulation of Si-SiO₂ defect generation in CMOS chips: From atomistic structure to chip failure rates," *IEDM Tech. Dig.*, pp. 93–96, 2000.
- [152] A. Lee, A. R. Brown, A. Asenov, and S. Roy, "Random telegraph noise simulation of decanano MOSFETs subject to atomic scale structure variation," *Superlattices and Microstructures*, vol. 34, pp. 293–300, 2003.
- [153] D. Vasileska and S. S. Ahmed, "Narrow-Width SOI Devices: The Role of Quantum-Mechanical Size Quantization Effects and Unintentional Doping on the Device Operation," *IEEE Trans. Elec. Dev.*, vol. 52, pp. 227–236, 2005.

